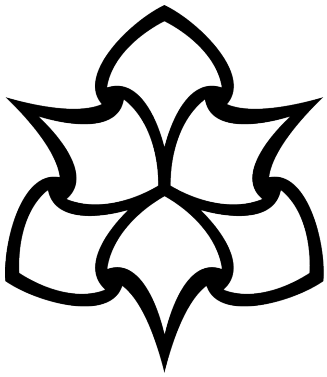


A framework for developing a conversational agent to improve normal age-associated memory loss and increase subjective wellbeing

Collette Curry

A thesis submitted in partial fulfilment of the requirements of Manchester Metropolitan University for the degree of Doctor of Philosophy.

School of Computing, Mathematics and Digital Technology,
Manchester Metropolitan University

2018

Abstract

Research has developed a baseline conversational agent (CA) framework that experiments suggest may improve normal ageing memory problems and increase Subjective Wellbeing (SWB) in participants aged 60+ with normal age-associated memory loss.

In 2008, 1.3 million people in the United Kingdom were aged 85+, this figure is projected to reach 3.3 million by 2033 (Morse, 2010). Thus, as the population profile changes, ageing memory impairment problems will become acuter (Morse, 2010). The number of people worldwide with diagnosed clinical memory problems is expected to double every 20 years to 66 million by 2030 and 115 million by 2050 (Casey et al., 2016, Prince et al., 2013). Improving memory impairment reduces distress for individuals and enhances wellbeing and independence (Dorin, 2007); (Wagner et al., 2010). The quality of life in old age can be improved by increasing SWB (George, 2010) that is concerned with how people experience the quality of their lives and includes both emotional reactions and cognitive judgments (George, 2010).

Experiments performed as part of the pilot study suggested evidence of increased SWB and improved memory after use of the CA. To support these early findings, modification to the agent and further experimentation was undertaken. Further work enhanced the preliminary work that was carried out and provided the opportunity to run further, more in-depth evaluations of the CA as both a reminiscence aid and as an improver of SWB.

This PhD study applied for and gained ethical approval (SE111219) from the Faculty of Science & Engineering Ethics Committee, Manchester Metropolitan University on 25 October 2012.

Copyright

Declaration

No part of this thesis has been submitted in support of an application for any other degree or qualification at this or any other institute of learning. Apart from those parts of the project containing citations to the work of others and apart from the assistance mentioned in the acknowledgments, this thesis is my own work.

Signed

_____

Collette Curry

Acknowledgements

There are many people that need thanking for helping me during my PhD studies. I thank all of them, they know who they are.

I wish to thank my exceptional supervisors, my Director of Studies, Dr. James O'Shea and second supervisor, Dr. Keeley Crockett, for their continuous support, patience, guidance, advice and encouragement during my PhD. studies. I appreciate their assistance greatly.

I would also like to thank Dr. Laura Brown, previously of Manchester Metropolitan University and now University of Manchester, for her advice during the early stages of my research. Her contribution proved invaluable to the ongoing development of my work.

To my dear friends, Tracey & Graham Dorber for your support and encouragement, I thank you most sincerely.

Most of all I would also like to thank my family, especially my husband Dr. Alan Curry for his support and encouragement. My parents for originally giving me the opportunities to succeed in education. I wish my father could have seen me complete this. To my children, James for always believing in me, and to Amy for her resounding enthusiasm and many words of encouragement whilst she undertook her own PhD journey. To my grandchildren, Amelia and Maggie, may this research encourage you to embark on your own journeys of educational discovery one day.

I'd like to acknowledge the help and guidance given by Alison Kershaw at St. Luke's Neighbourhood Centre and the members of the St. Luke's Art Project for their assistance, as well as encouragement during my research.

I dedicate this thesis to the memory of my dear friend Professor S. Barry Cooper (1943-2015), University of Leeds, whose unwavering support and encouragement helped me achieve my aims. His frequent checking in on my progress was much appreciated.

Table of Contents

List of Tables

LIST OF FIGURES

List of Abbreviations

ASR	Automatic Speech Recognizer
CA	Conversational Agent
CCQ	Communication Competence Questionnaire
DAML	DARPA Agent Markup Language
DARPA	Defense Advanced Research Projects Agency
DM	Dialogue Manager
DMS	Dialogue Management System
ECA	Embodied Conversational Agent
EMQ	Everyday Memory Questionnaire
EMQ-R	Everyday Memory Questionnaire Revised
GO-CA	Goal-Orientated Conversational Agent
HADS	Hospital Anxiety and Depression Scale
HCI	Human-Computer Interaction
NLP	Natural Language Processing
NM	Normal Ageing Memory
NLU	Natural Language Understanding Unit
OWL	Web Ontology Language
RDF	Resource Description Framework
SASSI	Subjective Assessment of Speech System Interfaces
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
SPANE	Scale of Positive and Negative Experience
SWB	Subjective Wellbeing
TBI	Traumatic Brain Injury
TTS	Text-To-Speech Engine
UMLS	Unified Medical Language System
UNSPSC	United Nations Standard Products & Services Code
W3C	World Wide Web Consortium
WIMP	Windows, Icons, Mouse, and Pointer
XML	eXtensible Markup Language

Chapter 1 Introduction

1.1 Overview

This introductory chapter defines the research questions and hypotheses of the research. The purpose of the Conversational Agent (CA) is to investigate its effect on normal ageing memory (NM) and Subjective Wellbeing (SWB). This research has developed a pilot reminiscing CA framework that may improve normal ageing memory and increase SWB in people aged 60+ with no medically diagnosed clinical memory impairment. The primary aim of the research undertaken was to investigate the design and use of a pilot reminiscing CA framework that may improve NM and promote SWB in people aged 60+ with no clinically diagnosed age-related memory decline. To achieve this aim the development of a CA framework into a baseline system for general use by this ageing population was produced. The research has identified that there were no existing CAs for reminiscing.

Therefore, the contributions of the CA include:

- The production of an ontology of reminiscence for inclusion within the proposed new CA framework. The use of a modular themed ontology is the first utilization of its kind for this class of CA.

- The development of a more efficient and productive scripting process using a WordNet model and utilizing an ontology of reminiscence. This led to improvements in maintainability.

- The establishment of short-term and long-term memory functionality. Enhanced CA engine and architecture, using a new scripting methodology aimed at solving novel problems in dealing with an ageing population, stimulating memory (storing and

recalling events for conversations) and enhancing wellbeing.

Initially, a pilot study allowed for the development of the CA framework and gathered findings about improved ageing memory and increased SWB amongst participants. Continued refinement and further testing gathered further evidence of positive effect suggested by the pilot phase. Usability of the system by an ageing population was also investigated to ensure maximum use of the system and to make any necessary adaptations. A hybrid (sometimes referred to as mixed-methods) research methodology was adopted. Such research has been practiced since the 1950's but formally began in the late 1980's and is becoming increasingly used by researchers (Creswell, 2003; Dunning et al, 2008; McKim, 2017). This combined qualitative and quantitative techniques, where a qualitative approach was utilised to uncover the meaning and attitudes behind quantitative data. Such research methodology is claimed by researchers to provide a more balanced perspective (Morse & Chung, 2003) and is therefore needed (Coyle & Williams, 2000). The qualitative phase included meetings with focus groups and informal interviews. The quantitative phase was done using questionnaires, experiments and conversational logs.

The research program contributed to knowledge regarding automated methods for teaching complex information to older adults with sometimes low computer literacy skills. Though over time, more people will have computer skills from their everyday working lives as they reach 60+. During the initial experiments, consideration was given for the usability of touchscreen input versus speech recognition for conversational interfaces. The acceptance and perception of a version of an Embodied Conversational Agent (ECA) by older adult participants were investigated. Specific areas of investigation included the use of verbal and nonverbal behaviour by older adults to indicate understanding or misunderstanding of verbally delivered content, and the prevalence, form, and function of gaze behaviour in a face-to-face conversation

with a reminiscing conversational agent. This work also made important contributions to the fields of human-computer interaction and embodied conversational agents, by providing information on the efficacy of different interface modalities and verbal and nonverbal communication strategies.

1.2 Research goal and objectives

- The research aimed to address the following research questions:

    1. Could interaction with a CA be a tool for improving normal age-related memory loss? Exposure to the reminiscence CA may improve access to existing memories.

        H_0: Use of CA will not make a significant difference in memory recall
        H_1: Use of CA will make a significant difference in memory recall

    2. Could the use of the conversational agent improve SWB?

        Exposure to the reminiscence CA may increase SWB.

        H_0: Use of the CA will not make a significant difference to SWB
        H_1: Use of the CA will make a significant difference to SWB

The research goal of this thesis is to develop a conversational agent reminiscing system that collects a knowledge base for future conversation with an individual during active conversation with a natural language interface. There are currently no conversational agents being used for reminiscence at this time.

The objectives of this research are:

1. To review conversational agents as an aid to reminiscence.

2. To develop a methodology for creating a conversational agent that dynamically learns during conversation with an individual. The CA has two

modes, firstly, long-term memory learn mode which saves the whole conversation and extracts keywords, matching to WordNet or the ontology of reminiscence. Keywords that do not match are saved collectively for manually adding to the ontology of reminiscence after being mapped to the system. The second mode is short-term memory only and recalls 25 iterations before replacing with new iterations as the conversation progresses. Learn mode can be activated remotely by the researcher and set to be available for a set period of time before reverting to short-term memory only.

3. To design an architecture for a conversational agent that engages users in reminiscing conversation.

4. To identify standard instruments to assess whether a conversational agent can be effective in improving normal ageing memory and subjective wellbeing.

5. To implement a prototype system and evaluate its success by conducting studies and running standard instrument questionnaires and statistical analysis of results gathered.

6. To develop experiences for the participants, to promote engagement.

7. To build up the storytelling potential of the CA. The CA has themed content including schooldays, the war years, the sixties. The CA engages the user in themed conversation retelling scripted stories of wartime experiences or experiences in the sixties. The music, the clothes, the cars are all included as themed stories as are wartime evacuation and food,

clothing and fuel rationing.

8. Evaluate the CA function and processing of natural language.

9. To investigate the feasibility of using a modular plug-in ontology to provide a dialogue system for the CA framework.

1.3 Detailed Contributions

The most significant contributions of this work are:

- Evidence that it is possible to engage people in conversation that allows for reminiscing and improves normal ageing memory recall. People using the system reported by using the EMQ-R questionnaire that they had better memory recall after sustained conversation with the CA than prior to conversation with the CA. The EMQ-R questionnaire is described in chapter 2.
- Evidence that natural language conversation with a CA that promotes reminiscing, increases subjective wellbeing. People using the CA completed the HADS questionnaire in the first experiments and a different group of people completed the SPANE questionnaire over the second experimental timescale. Results indicated increases in reported subjective wellbeing. The HADS and the SPANE questionnaires are described in chapter 2.
- Production of a modular ontology of reminiscence for inclusion within the pilot CA framework. The ontology of reminiscence was made up from data gathered from a reminiscence questionnaire and also from face-to-face discussions with the target group of participants. The reminiscence questionnaire is described in chapter 2.
- Utilization of an ontology of reminiscence within the CA system. This ensured that there was some themed conversation between the

participant and the CA. More reminiscence content was collected by the system and added to the ontology manually later.

- Implementation of short and long-term memory functionality within the CA framework. Short-term memory recalls the last 25 iterations during conversation with an individual. When the long-term memory is active during 'learn' mode, the CA remembers what has been said, it checks for keywords that map to the ontology of reminiscence or to WordNet, if the keyword is not found then the keyword is saved in an area of the database to be manually added to the ontology of reminiscence later. Automation of this process is possible but would require more rules to be created to check for abusive and other unwanted subject matter for conversation within the CA framework.
- Development of novel shorter scripting mechanism for the CA using the WordNet system as a model. WordNet is a lexicon of information comprising synsets. The synsets are made up of verbs, nouns, adjectives and adverbs in English, each expressing a distinct concept. Using a similar mapping technique, it was possible to create a method of including new categories and related keywords that the CA would understand. This resulted in less lines of scripting being required for every possible permutation of a word during conversation with the CA.
- Saving and referencing of content for future conversations with the same individuals, over time building up a personal knowledge base.

1.4 Features of the system

- Automatically performs spell-correction, idiom substitution, proper name and number joining.
- Topics about any subject can be implemented, linked and work within the modular system.
- Consists of rules organised into themed topics of conversation. Some rules match user input. Others lead the conversation on a particular subject. More rules react correctly to input and responses to conversation.

1.5 Thesis outline

The thesis is divided into seven chapters. Chapter 1 is an introduction and also outlines the research questions, hypotheses, contributions and research objectives. Chapter 2 looked at the literature on human memory, memory impairment, subjective wellbeing and the standard instruments to test for these factors. Reminiscence and triggers to memory retrieval was explored. The effects of ageing and the use of computer systems were considered and were later tested experimentally. Chapter 3 looked was dedicated to looking at CAs, their history, development and use as forms of natural communication between computers and humans. Different versions of CAs were discussed in this chapter and their production and use included. Development of an ontology of reminiscence was suggested and was implemented later in the thesis. Chapter 4 related the design and production of a prototype CA framework for reminiscence. An ontology production technique was defined and implemented. Chapter 5 described initial experiments with the CA framework. The use of the CA framework by older people was investigated to inform on the design process of the CA framework. The preference for a male or female entity was investigated and this informed the adoption of a female entity for the CA framework. Basic usability of the system was investigated, and results gathered and utilised in the development of the CA. Ease of use of standard instruments was noted and this informed future use. The conversational content was adapted and improved. Chapter 6 related a 14-day experimental period where subjective assessments took place with testing of various parameters including before, during and after exposure to the CA framework. The research questions were answered by the results and the results also indicated that the null hypotheses should be rejected. Chapter 7 summarises the research and contributions in relation to the research aims and objectives of this thesis. It also gives recommendations for the direction of future research.

Chapter 2 Human memory and standard instruments

2.1 Introduction

This section contains background on the elements of the research: human memory, memory impairment, SWB, reminiscence work and standard instruments for testing and evaluation.

2.2 Human memory

According to Fortey, 2008, people have a personal history of specific events from their past that they can draw upon to improve their learning and decision making. All our lives are collections curated through memory. We collect recollections and facts and store them, often half-forgotten or tucked away on shelves buried deep within the psyche (Fortey, 2008). This episodic memory was first described in depth by Tulving (Tulving, 2002, Tulving, 1983). Tulving's focus is phenomenological and distinguishes episodic memory from semantic memory. Episodic learning remembers events and history that are embedded in experience, while semantic learning extracts facts from their experiential context. Thus, a memory of looking out over the Egyptian pyramids during a previous family holiday is an episodic memory; however, if someone asks what condition the pyramids are in or their history, typically a person would use semantic memory to answer. The ability to remember where people have been, what they have sensed and what actions they have taken in various situations provides a knowledge base of information that is invaluable for acting in the present. It is as though a person can time travel within their memory. Knowing your personal history facilitates your ability to perform several cognitive capabilities in the context of sensing, reasoning, and learning:

Sensing:

• Noticing Significant Input – detecting what is important about a given situation by its relative familiarity

• Detecting Repetition – realizing when you are repeating the same series of actions and altering your behaviour as a result

• Virtual Sensing – retrieving past sensing of features outside current perception that is relevant to the current task

Reasoning:

• Action Modelling – predicting the immediate outcome of your actions

• Environmental Modelling – using experience to predict how the environment will change

• Recording Previous Successes/Failures – using past performance to guide future behaviour

• Managing Long Term Goals – keeping track of a plan and what steps in that plan have been accomplished so far

• Sense of Identity – understanding one's behaviour in relation to other agents

Learning:

• Retroactive Learning – reviewing experiences and learning from them when sufficient time (or another resource) becomes available

• Reanalysis gave new Knowledge – relearning from experience upon receiving new knowledge

• Explaining Behaviour – reviewing your past actions to others for mutual benefit

• "Boost" other Learning Mechanisms – provide a database of knowledge that can be manipulated by other learning mechanisms

Source: Tulving (Tulving, 1983)

Despite this array of benefits, the vast majority of conversational agents lack an episodic memory. This does not mean that most conversational agents lack all the cognitive capabilities listed above. Specialized algorithms can provide individual cognitive capabilities, but if these capabilities are to be supported, doing so with specialized algorithms would lead to redundant functionality that could be provided more efficiently by a single episodic memory. Therefore, a general, task independent episodic memory could be a summarily efficient approach to providing a conversational agent with a wealth of functionality. People encode, store and retrieve information from experience to create a retrospective memory. If it is not possible to encode information at the time the events are occurring, how can they be later recalled? (Howe and Courage, 2004). Memory from personal events occurs much later than first thought (Howe and Courage, 2004). Recalling experiences during early childhood is variable, depending on the child's age at the time of the event. Children at the age of 1-2 can recall personal events, though only in fragments when questioned several months later (Prabhakar and Hudson, 2018). Two-year-old children form autobiographical memories and remember them over periods of at least several months (Fivush, Gray & Fromhoff, 1987). Even if it's sometimes hard to recall experiences, they are permanently inscribed somewhere amid the billions of neurons in the brain (Furlow, 2001). This idea is not new; it is one that dates back to the early philosophers, such as Aristotle (Thorndike, 1905). A view restated at the start of the 20th century by Thorndike (1905) and similarly claimed by both Rank (Rank, 1924) and Freud in 1963 (Freud et al., 1995). There are several different kinds of memory that have been identified and memory impairment and episodic or retrospective memory research could be of relevance to the evolution of a reminiscing CA.

2.3 Memory Impairment

Studies of memory impairment in humans have been fundamental to learning about the organisation of memory and its cellular and

molecular basis. When memory impairment occurs, the question inevitably arises whether the impairment reflects impaired information storage or impaired accessibility (Bryant et al., 2005). A decline in functional memory is a normal part of the healthy ageing process, but levels of impairment are much greater in clinical conditions such as dementia. According to Bryant (Bryant et al., 2005), memory impairment can reflect the fact that there have been changes in the brain in the areas that represent storage. This is called storage deficit. Alternatively, it could represent a fault in the memory retrieval system. The memory might be intact, but the information is not accessible. This is called retrieval deficit. In the latter scenario, memories are possibly recoverable under different and appropriate circumstances. The use of a CA to prompt retrieval of retrospective memories will be the subject of experiments over the time of this project.

2.4 Subjective Wellbeing

Moods are subjective states of mind that are typically described and quantified using self- reporting measures (Brown and Astell, 2012). Mood states are also related to the broader psychological concepts of wellbeing and quality of life. Two distinct aspects of mental wellbeing have now been identified: subjective (hedonic) wellbeing and psychological (eudaimonic) wellbeing (Ryan and Deci, 2001). SWB is of the most relevance to mood states. A person who has a high level of satisfaction with their life, and who experiences a greater positive effect and little or less negative affect, would be deemed to have a high level of SWB [or in simpler terms, be very happy]. The concept of SWB falls within the hedonic perspective that defines wellbeing or happiness as being fundamentally about maximizing pleasure and avoiding or minimizing pain. This differs from the eudaimonic perspective, which, places focus on meaning in life and self-realization, and the extent to which a person fully integrates this into his or her life. When psychologists

measure SWB, they are measuring how people think and feel about their lives. The three components of SWB, life satisfaction, positive affect and negative affect, are independent factors that should be measured and studied separately (Andrews and Withey, 1976), (Lucas and Dyrenforth, 2006). Thus, the presence of positive affect does not mean the absence of negative affect and vice versa. Before evaluating the correlates and predictors of SWB, it is worth noting some of the instruments used in measuring the components of SWB. Life satisfaction can be measured using a questionnaire such as the 5 item satisfaction with life questionnaire (Diener et al., 1985), (Pavot and Diener, 1993). Affectivity can be measured by, for example, the PANAS [positive affect negative affect schedule] (Watson et al., 1988). Further instruments have been researched and several assessed and described in more detail in this report. Some of these instruments will be utilised to check SWB. The CA should stimulate SWB, and this will be subjectively measured using a relevant instrument after experimentation of conversation between the participants and the CA.

2.5 Effects of ageing

To create a tool adaptable for older people, it is crucial to consider the effects of ageing - Alzheimer's Association (2001, 2011). As people grow older, their abilities change (Heller et al, 2001). This change includes a decline in cognitive, physical and sensory functions, each of which declines at different rates relative to one another for each individual (Heller et al, 2001). This means that older people are not one consistent group and presents a challenge to system designers and programmers (Van de Watering, 2018). Older people often suffer from hearing and visual problems associated with ageing, as well as sensory problems caused by age-related illnesses. Visual problems are well-known (Jones and Van der Eerden, 2008), and the ability of

individuals to react to different types of visual and auditory prompting for assistance is important (Bouchard et al., 2012). There is often a gradual decrease in visual acuity including pattern recognition, visual search and processing, sensitivity to contrast, resistance to glare, weakened colour perception and light sensitivity. Decrease in hearing results in difficulties in hearing high frequencies and therefore when sound is used, it should be in the lower frequency ranges (Zhao, 2000). Older adults often suffer from indistinct speech, making the use of speech recognition for this age group problematic (Bouchard et al., 2012). These factors need to be considered when older people are to use computers. There are guidelines for designing websites given by various authors (Dickinson et al, 2003; Zhao, 2000). Recommendations on colours, fonts, navigation, sound, content, layout and style can be utilised during the design phase of the proposed CA framework.

2.6 Reminiscence

Triggers can be used to prompt retrieval of seemingly lost memories. Toseland (Toseland, 2005) uses the term reminiscence therapy. However, Gibson (Gibson, 2004) takes issue with this description, stating that reminiscence is not a therapy, which implies expertise and distance, but is better described as reminiscence work, which, Gibson claims, illustrates its participative nature. Reminiscence as a process is often considered universal (Butler, 1963), something we all do (Gillies and James, 1994). Since the Butler paper was published in 1963, there has been a significant increase in the literature concerning reminiscence and life review. Reminiscence concerns telling stories of the past, personal histories, individual perceptions of social groups inhabited, and events experienced personally or at a distance. Bender (Bender et al., 1999) provides a wide range of 20 possible purposes and benefits that can derive from reminiscence.

These include:

Purpose 1: Reminiscence to encourage spontaneous conversation between clients. People can get out of the habit of talking to other people because they have become physically or psychologically isolated.

Purpose 2: Reminiscence to encourage informal conversation. Getting clients to talk informally to each other and staff - in the canteen/over coffee/on the bus.

Purpose 3: Reminiscence as fun. An activity that is effective in helping you forget your problems and pains for an hour or two will be very welcome.

Other therapeutic benefits have been observed, such as increasing self-acceptance of ageing (Magee, 1994), promoting a holistic approach to care (Penn, 1994) as a counselling tool (Malde, 1988) as well as a social aid. Reminiscence is also associated with therapy and change but can be seen as a social and enjoyable activity without being construed as anything more (Bryant et al., 2005). If the CA is effective as a reminiscence companion, then similar therapeutic benefits from its usage will be evident. Older people like to reminisce or use narrative resources and think about their life. This exercise could be done around things that form part of their history of life and their sense of wellbeing. Neimeyer (Neimeyer, 2012) has developed a great deal of work on this, and one area he explores is in the area of bereavement and grief experience. Relating memories of bereaved people and their lives have a positive effect on people suffering from bereavement, but as Bender (Bender et al., 1999) pointed out, older people have a connection with reminiscence on many different levels. Such as with reminiscence materials such as photographs on the wall, objects that are readily touchable are all easy points around which to start a conversation. For people with age-related illnesses, especially Alzheimer's Disease, where the ability to form new memories or recall

recent ones is impaired, recalling older memories is often easier and very positive, helping to retain a sense of self and make sense of a world that, to the people concerned, is increasingly becoming incomprehensible. It also gives them something to think about and discuss. It is well established that keeping people with declining cognitive function mentally active helps them; reminiscing about what they can remember is an effective way to do this (Cohen and Taylor, 1998).

2.7 The Everyday Memory Questionnaire

It will be necessary to assess whether the CA has any effect on memory and recall by the participant. One disadvantage of everyday memory research is that, because of its diversity, most investigations have been of isolated phenomena such as eyewitness testimony (Loftus, 1996), autobiographical memory (Rubin et al., 1986), flashbulb memories (Conway, 1995), or oral tradition (Rubin, 1995). This fragmentation of the field (Gathercole and McCarthy, 1994) is not conducive to the identification of memory processes and structures that these phenomena might have in common (Cornish, 2000). In fact, Banaji (Banaji and Crowder, 1989) have argued that studying everyday memory phenomena can never identify general, underlying principles in the way that laboratory studies can. However, an approach which investigated a range of everyday memory phenomena simultaneously might provide us with the kind of evidence for common processes, if there are any, that is currently lacking, and provide a source of evidence about the human memory system which would be an alternative to laboratory-based investigations (Herrmann, 1984). Herrmann (Herrmann, 1982, Herrmann, 1984), has distinguished between questionnaires that directly test memory for facts or events (memory questionnaires) and those which rely on participants reports about their performance (metamemory questionnaires). Metamemory questionnaires have the advantage of greater scope and generality,

though their validity is more problematic (Banaji and Crowder, 1989), (Herrmann, 1982). They are essentially inventories of self-reports by people about their memories, i.e. of the beliefs people hold about their memories or of their recollections of their memory experience and performance. Most questionnaires which assess everyday memory performance are of the subjective or metamemory type (Gilewski and Zelinkski, 1986), (Herrmann, 1982), though a few present participants with a series of small memory tasks (e.g. the Rivermead Behavioural Memory Test (Wilson et al., 1986). Everyday memory questionnaires also overlap with more general tests of cognitive performance such as the Cognitive Failures Questionnaire (Broadbent et al., 1982). This includes material on motor skills, attention, and perception as well as on memory, though it has been treated as a memory questionnaire (Pollina et al., 1992). The Everyday Memory Questionnaire (EMQ) was developed as a subjective measure of memory failure in everyday life (Sunderland et al., 1996a). This questionnaire was initially designed for a study of everyday memory after Traumatic Brain Injury (TBI) (Sunderland et al., 1983) and was also used with healthy older adults living in the community (Sunderland et al., 1986). A modified version was used in a postal survey of TBI patients (Sunderland et al., 1984). Also, a shorter version was used in studies of memory after stroke (Tinson and Lincoln, 1987), (Sunderland et al., 1996b), and age-effects on normal memory (Sunderland et al., 1996a), [Appendix 6]. According to Sunderland (Sunderland et al., 1996a) this questionnaire was always intended as a research instrument and has not been published as a standardized assessment. Indeed, there are no norms from any large standardization sample (Hickox and Sunderland, 1992). The use of a revised shorter version of the EMQ, the EMQ-R [Appendix 7] both before and after exposure to the CA over time should suggest whether there has been any improvement in normal ageing memory after exposure to the CA.

2.8 The Hospital Anxiety and Depression Scale

To test Subjective Wellbeing there are several tests available, one such instrument that has been extensively used is the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983). The HADS is a self-reporting system, which is used as a marker for general wellbeing (Dunbar et al., 2000). The HADS is a 14-item scale (Zigmond and Snaith, 1983). Seven of the items relate to anxiety and seven relate to depression. It is a tool designed to detect anxiety and depression in people with reasonable physical health [Appendix 8] so may be particularly suitable for use with the participants initially testing the CA.

The anxiety questions include:

- I feel tense or wound up
- Worrying thoughts go through my mind
- I can sit at ease and feel relaxed

The depression questions include:

- I feel cheerful
- I look forward with enjoyment to things
- I feel as if I am slowed down

Each item in the questionnaire is scored on a response scale from 0-3 and this means that a person can score between 0 and 21 for either anxiety or depression (Bjelland et al, 2002). The HADS uses a scale and therefore the data returned is ordinal. Ordinal data is one of four possible scales of measurement, nominal, ordinal, interval and ratio. After adjusting for six items that are reverse scored, all responses are summed up to obtain two subscales. According to Zigmond & Snaith

(1983), recommended cut-off scores are 8-10 for doubtful cases and ≥11 for definite cases. The HADS has been recommended to assess psychological distress amongst a general population aged 65-80 years old (Djukanovic et al, 2017).

2.9 The Scale of Positive and Negative Experience Scale

Another such instrument for testing SWB is the Scale of Positive and Negative Experience (SPANE) (Diener et al., 2009). This scale is a 12-item questionnaire comprising six items for positive experience and feelings, and six items for negative experience and feelings. The two scores are combined to create an experience balance scale. This 12-item scale has a number of desirable features compared to earlier measures of positive and negative feelings. In particular, the scale assesses with a few items, a broad range of positive and negative experiences and feelings (Diener et al, 2009). The SPANE asks people to recall their activities and experiences during the past four weeks and to report these feelings. The scale can be used with different time frames, such as “Yesterday”, “The Past Week”, or “In general”. It is a very versatile tool (Diener et al, 2009). The SPANE scale includes the following options which are reported as ‘Very rarely, Rarely, Sometimes, Often or Very Often or Always’:

‘Positive, Happy, Good, Joyful, Pleasant, Contented, Negative, Sad, Afraid, Bad, Angry, Unpleasant’.

This instrument could inform the levels of SWB in participants before during and after conversing with the CA framework. It could be a better instrument for wider experimental use with groups of participants as it is not titled “Hospital Anxiety & Depression Scale’ and would therefore not conjure up connotations of being patients, or cause unease admitting to that a person is anxious or depressed like the HADS might engender away from a controlled medical environment.

2.10 Chapter highlights

- Standard instruments to test for NM ageing problems and SWB have been identified and selected for use within the research to be carried out.
- Human memory and related aspects have been investigated.
- Memory impairment studies have been looked at and these have been summarised.
- Subjective Wellbeing is normally assessed using self-reporting measures (Brown and Astell, 2012). Standard instruments have therefore been identified for using within the research being undertaken.
- The effects of ageing and the use of computer-based interfaces have been considered important to the research and will be tested experimentally.
- Triggers to retrieving memories and also reminiscence as a process we all do has been looked at.

Chapter 3 Conversational Agent Systems

3.1 Introduction

Natural language communication between a person and a computer has been a goal in artificial intelligence for many decades. This has resulted in the 'Turing Test' (Turing, 1950). The Turing Test was proposed to develop a machine that could imitate a human being well enough to fool a judge during conversation that it was really a person. There are three main types of Conversational Agents (CAs) that allow communication between computers and people using natural language. The first type of CA are known as chatbots and are text-based. They attempt to engage the user in intelligent conversation. Early attempts to pass the Turing Test were made by chatbots that used pre-programmed scripts to create an illusion of intelligence (Weizenbaum, 1966). There are more complex text-based CAs that solve specific problems in set domains. Other systems use spoken dialogue interfaces rather than text-based interfaces. Embodied CAs (ECA) can mimic human body language and gesture (Cassell, 2000).

CAs have been used to describe, text-based, spoken and embodied conversational systems (O'Shea et al., 2011). All CA systems must attempt to both understand the user input and provide an appropriate response to that input.

The goal and recent expansion of CAs in e-commerce, apart from being used to sell products and provide advice, is to provide a human with an answer to a question from an entity mimicking a human (Bogdanovych et al., 2005a). Today, CAs are being used in many different applications. Some examples are listed below:

- Virtual customer service assistants to aid online shoppers (Bogdanovych et al., 2005b).

- Information staff at airports (Bogdanovych et al., 2005a).

- Tour guides in museums (Michie et al., 2006).

- Living history guides for school children studying history (Curry et al., 2011).

- Proposed for student debt advice (O'Shea et al., 2010).

- Bullying and harassment advice (Latham et al., 2010).

As mentioned earlier, there are several types of CA available. Within CA systems, Dialogue Management Systems (DMS) are used to answer frequently asked questions in a general manner. They have rigidly scripted answers and do not engage in free natural conversation. An example of a DMS would be the British Telecom customer FAQ service or 'Lucy' on the 02-mobile phone company website. These attempt to cut back on the number of general inquiry emails being sent to their respective companies.

A Goal-Orientated Conversational Agent (GO-CA) is specifically designed to interact with a human, utilising natural language dialogue. Achieving a specific goal, typically a business goal, such as providing information or advice or selling a product or service. It plays the role of an empowered human in a productive application or task. Thus, the GO-CA may spend more time leading the conversation and asking questions than the human (O'Shea, 2010).

3.2 Embodied and Linguistic Agents

In addition to the types of CA identified, two forms can be identified: Embodied agents and Linguistic agents. Embodied agents possess an animated humanoid body and have facial expressions and eye movement (Cassell, 2000). Linguistic agents consist of spoken or written language without embodied communication (Cassell, 2000). A growing consensus within the human-computer interaction (HCI) community states that traditional WIMP (windows, icons, mouse, and

pointer) interfaces need to evolve and display a more adaptive, believable, flexible and human-oriented presence (Ball and Breese, 2000). In doing so, one of the main aims for HCI designers has been to develop a more natural human-computer interface (Holtgraves et al., 2007). A primary approach to addressing this has been a trend for making the interface more anthropomorphic and utilise CAs which behave in a manner that is more social. The initial drive towards creating an embodied, more anthropomorphic interface was based on early studies addressing social responses to text-based and talking face interfaces (Walker et al., 1994), (Sproull et al., 1996), (Khan, 2011). Walker (Walker et al., 1994) demonstrated that users responded more positively, spent more time, made fewer mistakes, and wrote more comments to a talking face than to a plain text display on the screen. Furthermore, the talking face was perceived as being more engaging than a text-based interface. Sproull (Sproull et al., 1996) concluded that users found the anthropomorphic interface emotionally more satisfying due to the interaction being regarded as easier, comfortable and more natural to use. Hence, such interfaces have the ability to make a computer more entertaining, engaging, approachable and understandable to the user, and make the user feel more relaxed with the computer (Catrambone et al., 2002), (Khan, 2011). Experimentation and further work will formulate the outcome of the project.

3.3 Development of conversational agents

After the world's first computer program that mimicked human conversation, Eliza (Weizenbaum, 1966), further enhancements in the design of such programs led to PARRY (Colby, 1975). PARRY was a paranoid, self-obsessed character capable of exhibiting personality, character and paranoid behaviour and who did not have to be a good listener. PARRY possessed a large collection of conversational tricks, including admitting ignorance in response to questions (O'Shea et al.,

2011). With such programs or chatbots as they are called, the primary aim has been to fool the user into thinking that the program's output has been produced by a human (the Turing test). Chatbots, reform the input and then output questions based on this input (Curry et al., 2011). Most chatbot scripts use a pattern matching technique, which has been identified as one of the most common and capable methods for developing chatbot dialogues that seem to be coherent and intelligent to users (Bickmore and Giorgino, 2006). Some chatbots use sophisticated natural language processing systems, many scanning for keywords within the input and pulling a reply with the most matching keywords, or the most similar wording pattern, from a textual database (Mauldin, 1994). The conversation follows the same structure each time the startup parameters is met, so the experience for individual users is not unique. Many chatbots now include functional features such as games and web searching abilities as well as conversation (Mauldin, 1994, Curry et al., 2011, Bickmore et al., 2013).

SitePal (SitePal, 2012) a company developed after over four years of research at Stanford University have an AI feature that allows a chatbot avatar web interface to be produced using Artificial Intelligence Mark-up Language (AIML): the simple open standard language. AIML was designed by Dr. Richard S Wallace. AIML is an XML-based language and relatively easy for anyone to learn. AIML was originally used to create a chatbot called A.L.I.C.E (Artificial Linguistic Internet Computer Entity) which won many awards, including the Loebner Prize on several occasions (Wallace, 2009). Despite AIML's deliberate simplicity, the language powers the most complex chatbots on the market to date, including the multi award-winning Mitsuku chatbot (Wallace, 2016). Indeed, Mitsuku created by Steve Worswick was awarded the Loebner Prize for the fourth time in September 2018.

Chatscript is a 'next generation' chatbot engine that powered a chatbot named Suzette, that won the 2010 Loebner Prize competition. Many

chatbots have now been programmed using this system. Chatscript aims to pattern match on general meaning using a simple visual syntax (Wilcox, 2011). Chatbots have developed much better conversational capabilities over time. They are however, totally scripted and do not interpret statements and then answer them intelligently.

More advanced CAs are also software programs, which differ from chatbots in that they do interpret as well as respond to statements made by users in ordinary natural language rather than reflecting the statement made by the user as a chatbot would. They integrate computational linguistics techniques with communication normally over the internet (Cassell, 2000). Conversational Agents are not necessarily web-based though. They are also present on various platforms, such as mobile phones, personal digital assistants, and other mobile devices.

Another newer method of CA production is the AI method (O'Shea, 2010, O'Shea et al., 2010) which looks at and compares semantic similarity of phrases (Li et al., 2003). Research is still relatively new in this area, but it is thought that such AI method systems will reduce time and effort in scripting CAs. This PhD research also aims to reduce time and effort in scripting CAs by exploring the use of modular ontologies within a CA framework.

ECAs are computer-generated, often cartoon-like characters that demonstrate many of the same properties as humans in face-to-face conversation, including the ability to produce and respond to verbal and nonverbal communication (Cassell, 2000). They constitute a type of (a) multimodal interface where the modalities are those natural to the human conversation: speech, facial displays, hand gestures, and body stance. (b) Software agent, insofar as they represent the computer in an interaction with a human or represent their human users in a computational environment (as avatars, for example). (c) Dialogue system where both verbal and nonverbal devices advance and regulate the dialogue between the user and the computer. With an ECA, the visual dimension of interacting with an animated character on a screen plays an intrinsic role. Not just pretty pictures, the graphics display visual features of conversation in the same way that the

face and hands do in face-to-face conversation among humans (Cassell, 2000).

Some major ECAs developed to date are 'Steve' (Rickel et al., 2000). Also, a *'How was your day?'* prototype companion system (Cavazza et al., 2010), and 'Olga' (Beskow and McGlashan, 1997), as well as 'Gandalf' (Cassell and Vilhjalmsson, 1999), and various other pedagogical agents. There is also a growing number of commercial ECAs, such as those developed by Extempo, Head pedal, and Artificial Life, and the Ananova newscaster developed by Ananova, Ltd. A social interactive companion robot called 'Mario' allows reminiscence of past holidays through the use of old family photographs and pre-programmed memory information to allow some reminiscence interaction with the user (Casey et al., 2016). 'Woebot' is a fully automated conversational agent system devised to deliver Cognitive Behavioural Therapy (CBT) to young adults with symptoms of anxiety and depression and successfully reduced symptoms of depression and anxiety and increased wellbeing and mood amongst a sampling of young people. Based on the 'Woebot' results it demonstrates that CA's appear to be a feasible, engaging, and effective tool to deliver CBT (Fitzpatrick et al., 2017).

These systems mentioned vary widely in their linguistic capabilities, input modalities (most are mouse input/text input/speech output only), and task domains, but all share the common feature that they attempt to engage the user in natural, full-bodied (in some sense) conversation.

There are several reasons why ECAs could provide an effective medium targeted at older adults with normal ageing memory loss. Firstly, the Human-Computer Interface relies only minimally on text comprehension (or not at all, in the case of speech recognition) and uses the universally understood format of a face-to-face conversation,

thus making it less intimidating and more accessible for a variety of people. Also, one study using a pedagogical agent, *'How was your day?'* (Cavazza et al., 2010) compared information delivery to students via an ECA that used speech output with an identical system that used text output instead. Finding that students recalled more in the speech condition than the text-only output (Cavazza et al., 2010). Functional older adult computer literacy is a relatively new field of inquiry, but one with potentially significant impacts on personal health, especially for older adults, minorities, and people with low levels of education or socio-economic status. The studies described in this thesis will provide valuable information about how comprehension of the conversational agent can be maximized in older people, as well as additional data on the prevalence of functional computer literacy problems and their impact on SWB and memory-knowledge outcomes. The work carried out for this research is to create a pilot reminiscing CA framework with domain specific components like an ontology of reminiscence, to boost the knowledge base. There is a requirement for the system to allow for user interaction, saving and referencing of content for future conversations with the same individuals, over time building up a personal knowledge base.

3.4 Unhappiness and negativity during conversations with a CA

The pilot reminiscing CA framework will direct people to seek help from external bodies if unhappiness and negative feelings are predominant in the conversation. The pilot reminiscing CA framework cannot take on the responsibility for helping with these issues but can direct a person to seek professional intervention. Indeed, information about how to contact a free, confidential '24-hour' helpline for older people, called 'The Silver Line' has been included in the conversational content. One of the pilot reminiscing CAs roles would be to help individuals feel better about themselves and thereby feel more

valued and increase their SWB. The conversations undertaken with the pilot reminiscing CA framework would re-enforce this role.

3.5 Conversational agents scripting

Most CA scripts are organized into contexts consisting of some hierarchically organized rules. CA's rely on scripts consisting sets of rules combining patterns and associated responses. This requires inclusion of a large library of rehearsed responses. More complex CA's that deal with a wide variety of inputs need many rules, each of which deals with an input pattern and the possible variations (Sammut, 2001). Each rule consists of a list of structural patterns of sentences and an associated response. User input is then matched against the patterns and a suitable response given by the CA (O'Shea et al., 2011). InfoChat is one such pattern matching system. Pattern matching systems scripting of CAs is very labour intensive and requires lines and lines of script, rules, and code. It is proposed to use an ontology of reminiscence to inform the knowledge base and thereby cut down on the script required. If the ontology was derived from oral interviews, written questionnaires and virtual conversations with a pilot group then the CA will become themed to the target group and provide an ontology of reminiscence for the CA. CA systems contain the following phases: **Source: (Jurafsky and Martin, 2009).**

1. The user speaks, and the input is converted to plain text by the system's input recognizer/decoder, which may include:
    - Automatic speech recognizer (ASR)
    - Gesture recognizer
    - Handwriting recognizer

2. The text is analysed by a Natural language understanding unit

(NLU), which may

include:

- Proper Name identification
- Part of speech tagging
- Syntactic/semantic parser

3. The semantic information is analysed by the dialogue manager that keeps the history and state of the dialogue and manages the general flow of the conversation.

4. Usually, the dialogue manager contacts one or more task managers that have knowledge of the specific task domain.

5. The dialogue manager produces output using an output generator, which may include:

- natural language generator
- gesture generator
- layout engine

6. Finally, the output is rendered using an output renderer, which may include:

A text-to-speech engine (TTS) or a talking-head avatar or robot.

CA systems that are based on a text-only interface contain only stages 2 to 5 (Jurafsky and Martin, 2009).

The InfoChat system is a commercially available pattern-matched goal-orientated CA (GO-CA) marketed by ConvAgent Ltd, a spinoff company at Manchester Metropolitan University that provides business rule automation with natural language interfaces (Convagent, 2005). InfoChat uses a sophisticated PatternScript scripting language (Michie and Sammut, 2000). Rules consisting of stimulus patterns and responses make up the scripts (Michie and Sammut, 2000). Each matched pattern generates a response to any input (Latham, 2011). InfoChat was used as an online advice system with a CA called Adam, a student debt advisor (Crockett et al., 2009), it was also used as an expert system to support bullying and harassment policies (Latham et al., 2010). The InfoChat system was investigated and some features such as the ability to have a number of patterns within a single rule, and rules organised into contexts for conversation were utilised within the pilot reminiscence CA.

InfoChat consisted of the following layers before modification for the pilot reminiscence framework.

An Aliza Layer - This is like an Eliza layer and contains general chat. It allowed the conversation to be sustained and redirected when no goal-oriented content in user utterance. It allowed the CA to respond to social remarks included in the discussion. It also included light-hearted responses to personal, challenging or antisocial comments by the user.

An Oracular Layer - responsible for answering questions put to the system. These include the many possible requests for information or clarification required to supply the attribute values to the CA, as well as general questions.

Conventional dialogue - This put questions to the user to acquire attributes, perform clarification tasks and provide the instructions

(diagnosis).

Filter -This is executed every time the user types an utterance, regardless of the current context in the main dialogue. It performs two tasks. Firstly, it contains a small number of rules to detect highly obvious statements. So, if a user volunteers additional attributes in the current context, they will still be recognized and captured. Secondly, it is used to detect racist or other highly offensive language in the conversation -which results in the communication being terminated (O'Shea et al., 2011).

The InfoChat system was already set up in the following way: A user utterance is first passed through the Filter and then (if the desired attribute is not captured) to the Conventional Dialogue layer. If the conventional dialogue layer does not have rules that can process the utterance, it is analysed to see if it is a question, if so the utterance is passed to the Oracular layer, if not it is passed to the Aliza layer. The InfoChat system has one simulated emotion, irritation. This would build up or dissipate over some dialogue transactions. If a certain level was reached the InfoChat system terminated the conversation. Happening very quickly (in the case of extremely offensive language from the user) or more gradually with warnings - in the event of the user failing to co-operate with the CAs conversational strategy (O'Shea et al., 2011).

3.6 Evaluating conversational agent systems

Evaluation of CAs is often subjective by their nature, evaluation of such systems is mainly done either by distributing a questionnaire to the users trying to reveal their subjective assessment of using the CA system or by studying the resulting dialogue from the conversational logs. The PARADISE framework (Walker et al., 1997) is one technique of CA evaluation. Other evaluation systems have been

developed and are useful as markers of conversational dialogue quality (Robinson et al., 2010). To compare the subjective and the objective measures a similar correlation study could be conducted, where values for subjective and objective metrics for both the whole group as well as any sub-groups could be compared. Also, questionnaires based on Subjective Assessment of Speech System Interfaces-SASSI; (Hone and Graham, 2000) and the Communication Competence Questionnaire-CCQ; (Monge et al., 1982) are sometimes utilised to test the effectiveness of CAs during natural language conversation with participants. Subjective assessment of the pilot reminiscence CA will be carried out.

Initial evaluation of the reminiscence CA will be undertaken using the HADS, self-reporting system that will test SWB (Dunbar et al., 2000) of the person both before and after application of the CA [Appendix 7]. The HADS is an established instrument used widely within the health service and is simple to administer. Also, the use of the EMQ (Sunderland et al., 1983) both before and after the application of the CA will inform whether there is a noticeable difference in NM after exposure to the CA [Appendix 6]. As related earlier, the EMQ is used widely as a subjective measure of memory failure in everyday life. This more direct assessment of the errors experienced by older adults during their daily activities may be more useful for directing the research into developing an intervention that will have a practical impact (Ossher et al., 2013). Wellbeing can be tested with a range of mood assessment techniques, including self-reporting measures (Ossher et al., 2013). The SPANE questionnaire could be administered to users of the pilot CA framework. The results could be collected, and data analysed to show levels of satisfaction with the reminiscing framework CA that could result in high levels of SWB.

3.7 Ontology development

The use of ontologies in computer science has been steadily

emerging in the discipline over several decades. The evolution of the semantic web has encouraged the development of ontologies. An ontology represents the shared understanding and the well-defined meaning of a domain of interest, thereby enabling computers and people to collaborate better (Wagner et al., 2010). Before going further into detail, an important step is to define the notion of ontology. Very briefly, an ontology is a formal specification of a conceptualization (here a domain), and it includes the definition of classes, objects, properties, relationships, and axioms. Ontologies are expressed using a formal language such as RDF or OWL and automatic support inference. Ontologies imply a kind of consensus within a community, meaning that they formalize concepts accepted within this community. There are many kinds of ontologies such as upper-level ontologies, task ontologies, and domain ontologies. Here, we were particularly interested in domain ontologies.

There are many approaches available to creating a domain ontology. No one method can be said to be correct or incorrect. The use of an ontology of reminiscence with the CA suggested will allow the sharing of a resource of remembrance and past times. Providing a new source of conversational material to be shared and explored with family members and friends as new conversational topics for person-to-person conversation after exposure to the CA. Review of the literature has found no systematic use of ontologies in the design of CAs to date. Ontologies have become common on the World-Wide Web. The ontologies on the Web range from large taxonomies categorizing Web sites (such as on Yahoo!) to categorizations of products for sale and their features (such as on www.amazon.com) (Noy and McGuinness, 2011). The WWW Consortium (W3C) is developing the Resource Description Framework-RDF; (Brickley and Guha, 1999), a language for encoding knowledge on the Web to make it understandable to electronic agents searching for information. The Defense Advanced Research Projects Agency (DARPA), in conjunction with the W3C, is developing

DARPA Agent Markup Language (DAML) by extending RDF with more expressive constructs aimed at facilitating agent interaction on the Web (Hendler and McGuinness, 2000). Many disciplines now develop standardized ontologies that domain experts can use to share and annotate information in their fields. Medicine, for example, has produced large, standardized, structured vocabularies such as SNOMED (Price and Spackman, 2000) and the semantic network of the Unified Medical Language System (Humphreys and Lindberg, 1993). Broad general-purpose ontologies are emerging as well. For example, the United Nations Development Program developed the UNSPSC ontology, which provides terminology for products and services (Noy and McGuinness, 2011). Ontology learning involves the use of some data (structured or semi-structured or unstructured) as input to the learning process. Structured data refer to already defined knowledge models including database schemas or existing ontologies. Semi-structured data designates the use of some mixed structured data with free text such as Web pages, Wikipedia, dictionaries and XML documents. Unstructured data is related to any textual content. The existence of structure helps direct the ontology learning process towards relevant parts of data.

Most of the approaches for ontology learning from (semi) structured sources rely on linguistic and statistical techniques and use the underlying schema already available in the structure. For instance, some works rely on dictionaries such as WordNet and try to parse natural language definitions. Examples of such works include OntoLearn (Navigli et al., 2004, Velardi et al., 2005) and (Rigau et al., 1998). Others rely on thesauri as the knowledge source (Van Assem et al., 2004) or on XML schemas (Volz et al., 2003), which get converted into domain ontology by translating non-terminal and terminal symbols into concepts and roles. Similarly, the work of (Stojanovic et al., 2002) uses a rule mapping scheme to convert an XML schema or a relational database schema into a domain ontology. The work of Deitel (Deitel et al., 2001) created an ontology learning procedure from RDF

annotations and (Nyulas et al., 2007) developed a plug-in for importing relational databases into an ontology editing environment (Protégé). Rule Knowledge bases have also been used to create an ontology (Suryanto and Compton, 2001) using statistical measures. The work of (Jannink and Wiederhold, 1999) extracted a graph structure from dictionaries and used statistical filtering and the PageRank algorithm to determine meaningful relationships and concepts. Another example is the work of (Papatheodorou et al., 2002), which built taxonomies using cluster mining from XML or RDF domain repositories.

In general, knowing the knowledge source structure helps and guides the ontology learning process. However, much of the available electronic data is in the form of unstructured documents. Being able to exploit vast domain corpora and electronic publications is then a requirement for ontology acquisition. WordNet can be used to extract synonyms, antonyms and other kinds of relationships. Involving also the detection of the right sense of the term and thus the use of word sense disambiguation algorithms (Papatheodorou et al., 2002).

3.8 Concepts relating to CA research.

Ontology – The concept of producing a reminiscence themed ontology that could fit into a specially designed CA framework has been considered. Linking such an ontology to WordNet is possible and would make the CA framework a more robust conversational system.

WordNet – An existing lexicon that has been mapped to successfully by such systems as the four times Loebner prize winning chatbot system of Bruce Wilcox. Mapping the reminiscence CA framework to WordNet would provide stable content for the conversation and allow for shorter scripting due to the existing synsets within WordNet that could be mapped to the CA framework.

Eliza – CA systems often use an Eliza layer to allow sustained conversation and to redirect the conversation when there is no goal-

oriented conversation from the user, thereby getting any themed conversation back on track. The student debt advisor, Adam, using the InfoChat system had an Aliza layer that did this. Adam's Aliza layer also gave responses to personal, challenging or antisocial comments uttered during conversation. Using an Eliza type layer within the design of a CA serves many important services. Providing more depth to any conversation.

Rules – Within CA systems rules are set to process and direct the conversational outcomes and to deal with user utterances. If a rule is met then there is a predetermined outcome. Highly obvious statements are detected, and unexpected statements can be dealt with to offer cohesive conversation.

Pattern matching – CA systems use pattern matching to classify the user utterances and therefore produce a suitable response for the person talking to the system. Chatbots use scripted content that allows the chatbot to respond to anything relating to the pattern. The system would then give a correct output. CA systems use advanced algorithms so for each kind of question asked by a user, a unique pattern must be available in the database to provide a correct response to the user.

Spell-checking – The inclusion of a spell-checking feature within a CA system would allow the system to correctly interpret what has been said by the user and therefore give an appropriate response.

3.9 Conclusion

This chapter introduced CAs, which allow humans to communicate with a computer interface using natural language. Some varying types of CAs were described and the development of CAs discussed. The development of CA scripts is very labour intensive and a time-consuming task. This could have a negative impact on commercial development costs of producing a CA for some applications. This PhD research will attempt to design a modular system that will utilise a plug-

in themed ontology of reminiscence as its knowledge base and cut down on the intensive scripting that is normally required. The possible content of user utterances can be anticipated as the pilot CA is themed towards a reminiscence domain. The use of the pilot CA as a web interface will be tested on small numbers of people, so handling the required number of conversations simultaneously in real time should not cause a problem. Scalability will need to be built into the system once the pilot stage is completed with sophisticated CA approaches, like natural language processing (NLP) (O'Shea et al., 2011) if the CA is to be made available to a larger audience. This may form future development work of the reminiscence CA system. NLP is a branch of artificial intelligence that helps computers understand human language as it is spoken, by interpreting and then processing human language.

This research has so far identified an unexplored area for research on using a CA as a conduit for reminiscence. Improvement in normal ageing memory (NM) as well as increased SWB should result if the research is successful.

3.10 Chapter highlights

- CAs allow natural language communication between computer interfaces and human users.
- The majority of text-based CAs use pattern-matching systems. They are fast enough to respond to user utterances in real time.
- There are different versions of CA within the field of Artificial Intelligence. Some were included within this chapter.
- CA script production is labour intensive, time consuming and often complex. It results in long scripts and lots of variations to expected user conversation must be scripted.
- TTS can be utilised to enhance the user experience but usability for older adults needs to be considered.

- There has been an expansion of CAs within e-commerce, also CAs are being used in a multitude of different applications. Some were described in the chapter.
- The chapter included CAs development from the first interface that mimicked human conversation to embodied conversational agents in use today. It also talked about evaluating CAs, which is often subjective.
- InfoChat is a successful text-based pattern matching system.
- Development of an ontology of reminiscence was suggested so that themed ontologies could be used within the CA knowledge base as modular plug-ins.

Chapter 4 Prototype of the CA Framework

4.1 Introduction

The primary aim of the research is to investigate the use of a pilot reminiscing CA framework that improves normal ageing memory (NM) and promotes increased SWB in people aged 60+ with normal non-clinically diagnosed age-related memory decline. To investigate, it was necessary for the development of a CA framework into a baseline system for general use by an ageing population. The pilot study allowed for the development of such an entity and gathered initial findings of improved NM and increased SWB amongst participants. Continued refinement and further testing will be necessary to gather further evidence of any positive effect suggested by the pilot phase. Usability of the system by an ageing population can also be investigated to ensure maximum use of the system and to make any necessary adaptations. The proposed research program will contribute to knowledge regarding automated methods for teaching complex information to older adults with sometimes little computer literacy skills. Though over time, people will have computer skills as they reach 60+ from their everyday working lives. Information was obtained on the usability of touchscreen input vs. speech recognition for conversational interfaces. The acceptance and perception of a version of an Embodied Conversational Agent (ECA) by older adult participants were investigated. Specific areas of investigation included the use of verbal and nonverbal behaviour by older adults to indicate understanding or misunderstanding of verbally delivered content, and the prevalence, form, and function of gaze behaviour in a face-to-face conversation with a reminiscing conversational agent. This work will also make important contributions to the fields of human-computer interaction and embodied conversational agents, by providing information on the efficacy of different interface modalities and verbal and nonverbal communication strategies.

Before the CA system was developed a number of systems were examined to see how they worked, these were described in the previous chapter. A suitable methodology was followed to create the system. The first factor to consider was 'What is the CA system going to offer?'. After considering this and the scope of the system it was decided that the CA system would provide information, answer questions, provide companionship and discuss different domain themed topics. The direction of the interaction would be dual, so that the user or the CA can introduce topics into the conversation. The CA would maintain the conversation through an Eliza layer and new content would be added to the personal knowledge base for each user. This would allow the user to modify the system to suit their reminiscence experiences. The CA would allow some 'small-talk' but would steer the conversation when required. It was hoped that this would build up rapport with the CA system. The handling of unknown or unexpected user input needed to be addressed and a plan to handle such content was built into the system. This was to ensure that appropriate CA responses were given at all times. This also allowed errors to be captured and to be handled by the CA and conversation continued uninterrupted. Simple pattern-matching wouldn't be enough to cope with unrecognised phrases, alternative grammar or text slang input. Therefore, a wide range of input variations, including misspelt words had to be recognised by the system. These variations were assisted by the inclusion of the WordNet system as well as a reminiscence ontology. The initial evaluation of the CA system provided conversational logs that were examined to see how this approach worked.

Of all the CA systems that were examined, none represented a reminiscence domain, and therefore this was to be a new reminiscence themed domain for a CA system. It was decided to utilise facets of existing CA systems as well as creating new functionality for the reminiscing CA framework. It was planned that the user would log in to access their personal knowledge base, which would build up over time, and to access a 'Learn' mode which would be set to run for either a set

period of time during experimental studies, or until the user ended the conversation. The Eliza layer would provide a conduit for the initial conversation and also keep the conversation going should the user go off track or lose interest in the current topic of conversation. A rule-based system would pass the input to the correct next layer in the system and an appropriate response would be given. A modular approach was adopted in terms of themed ontology inclusion. A framework was created into which new content could be added and utilised during conversation. Content for the themed ontology was gathered during group discussions, use of a questionnaire and also from the conversational logs. A specification for the software requirements was drawn up and is included in the following section.

4.2 Software Requirements Specification

The aim of the research was to create a reminiscence CA framework that would engage older people in conversation through reminiscence. The purpose was to improve ageing memory and increase subjective wellbeing. The intended audience was older people aged 60+ who have no clinically diagnosed memory problems. Intended use was to act as a companion system to older people who could potentially converse with the system at any time of the day or night. The scope of the study was for older people to converse freely with the CA system and to have their conversations revisited whilst talking to the system. Benefits of talking with the CA include being able to share conversation with a non-judgemental, never tiring, always available system. Producing the CA system will allow for greater understanding of what is possible within a software building environment for the developer. There is the potential for a CA system to upset the person talking to the system by having them recall something from their past that is upsetting. The system cannot act as a counsellor or as a medical professional, however, it should have a

safety net built in to direct the person speaking to the system to a source of help.

Betty is a new CA system that promotes and encourages reminiscence with older people who may have some normal ageing memory decline but who do not have a clinical diagnosis of memory loss.

Primary users are aged 60+ and it is thought that they would use the system on a regular basis. Several factors were also considered:

Usability - whether the system could be used by older people,

Access to computer interfaces for older people,

Relevant initial conversational content was incorporated into the system to promote ongoing conversation.

There are assumptions that all people can use computer technology and that people will feel happy about doing so. Research into older people using such technology has indicated that given time older people can adapt and use systems efficiently but that their needs need to be considered almost on an individual basis.

The end user and their ability to use the system was considered and experiments carried out in chapter 5. The hardware available to the target group was considered and experiments carried out on different systems, included in chapter 5. Other non-functional requirements were considered and included performance, safety, security and quality. Performance was planned to give a response rate by the CA system of no longer than 5-seconds. The system needed to work efficiently on different computer interfaces, from ‘chrome-book’ to ‘Windows 10’ to ‘Apple’ computers. Safety of the system in terms of rekindling bad memories was considered and advice given on where to seek help. The potential bad memory would not want to be restated by the system so once such a bad memory was detected by the system and seeking help advice given, then the bad memory would be locked so that future conversation wouldn’t mention it. The CA system had a login requirement that helped protect against unauthorised access to the conversational system. The quality of the system was informed by experiments carried out in chapter 5 and chapter 6. Conversational content was increased

after a number of conversational iterations with the people taking part in the studies.

Following on from the introductory section and the methodology adopted for design of the reminiscence themed CA system, these are examples of the components designed using this methodology:
Conversation Manager, Eliza layer, Rule Matcher, WordNet, Reminiscence Ontology, Spell-checker. It also had the following functions: Short-term and Long-term memory.

These components and functions are described below:

Conversation Manager

The purpose of this is to accept user input and pass it to the Eliza layer or through the Eliza layer and direct to the rule matcher.
It was designed to be able to accept user input that is expected or content that is unexpected and pass it through to the Eliza layer or direct to the rule matcher. To use the CA system, a participant visits the CA framework webpage and is greeted with a login page. Once the user has signed in successfully the conversation manager saves the user input and passes it to the rule matcher for processing. If a user has been in conversation previously then the rule matcher will activate immediately. If they haven't spoken previously or decide to not continue with a previous conversation then the next historic conversational topic will be invoked and if refused then the Eliza layer will be activated to begin to guide the conversation. It was necessary to evaluate different user interfaces to allow a user to enter their name and ongoing conversational content. This was to ensure that the majority of older people taking part in the study could use the interface successfully.

Eliza Layer

The function of the Eliza layer is to keep the conversation on track and to offer up general conversational content for discussion with a user. The

Eliza layer comprises more than 14,000 conversational items. These were scripted to provide general conversation as well as to guide the conversation. Conversation can be sustained and continued generally or redirected to the Rule matcher if known content is encountered. Unknown content is passed into a log file for manually adding to the Reminiscence ontology. The Eliza layer allowed up to 20 hours of general conversation. The system also retains the last 25 user iterations, and these are reused within the general conversation. These are saved in the user text file within short-term memory. Conversational logs indicated that the general conversation could be maintained for more than several hours. When the system is set to general mode this conversation is enhanced by the inclusion of data from the last 25 user iterations. When the system is set to 'Learn' mode, the data is sent via the rule matcher to determine if the content fires rules and leads the conversation into the reminiscence themes, or if the rules don't fire because content is unknown then the content is passed to the log file for manually adding to the ontology later. Failing to find an existing theme or context will return conversation combining information from within the past 25 user iterations and general conversation from the Eliza layer.

Rule Matcher

The rule matcher accept suser input and checks that it matches known data then return an appropriate response. It consists of different topics and concepts which match to give appropriate responses to user utterances. Rules are set to allow them to fire in order of most likely appropriate responses. The rule matcher comprised more than 20,000 rules which mapped to WordNet and the Reminiscence ontology. The conversational logs indicated which rules fired and when. They also indicated the paths conversations followed and also collected new content for inclusion within the Reminiscence ontology. A 5-point Likert scale assessed subjectively whether the CA system responded well during conversation with people taking part in the experiments.

Competency questions tested the system for accuracy of expected responses.

WordNet

WordNet is a widely used resource in natural language processing. It is a lexicon database in the English language consisting of groups of cognitive synonyms or synsets, each expressing a distinct concept. These synsets are interlinked via conceptual semantic and lexical relations. This database can provide content for the CA framework if the framework is mapped to it successfully. WordNet has 117,000 available synsets which link to other synsets using conceptual relations. Each synset holds a brief definition and short sentences illustrating the use of the synset members. Identification of the concepts within WordNet and converting the WordNet system to text files sorted from a - z and numbered text files for dates allowed the CA framework to be mapped to them. Looking for missing synsets within WordNet informed the construction of a themed ontology of Reminiscence. WordNet has certain biases for synsets that don't cover every item identified as topical for reminiscence. For example, 'coalman' is included in WordNet, but not 'coal hole'. WordNet contains common nouns, verbs, adverbs and adjectives. WordNet is invoked within the CA framework by calling the keyword and then the index of the meaning required is returned. For example, user enters "I remember the coalman', system takes in keyword '~coalman'. This searches the c text file and finds a definition about a coalman being a person who delivers coal, this would be returned by the CA to the conversation. In addition, the c file within the reminiscence ontology will be searched and content found prepared for further conversation with the user. Not enough reminiscence themed content was included in WordNet, so an ontology of reminiscence was required to supplement this. Conversion of the data into text files allowed the database to be accessible within the CA framework.

Reminiscence Ontology

The reminiscence ontology provides content relevant to the themed domain. The design of the reminiscence ontology followed an ontology production technique (described in section 4.5). The aim was to create a collection of reminiscence themed conversational content that would promote recall and continued engagement with the CA. A collection of files, sorted from A-Z and with number and date files was added to the CA framework. These included data collected from face-face conversations, a reminiscence questionnaire and from CA system identified unknown keywords collected during CA conversation. This became the ontology of reminiscence. The type of subject that people commonly like to reminisce about became apparent from the conversational logs. Some areas of the ontology were never called upon, such as 'The Festival of Britain', others such as 'classic cars of the Sixties', 'Music from the forties', 'Schooldays' were often spoken about during the experiments. Feedback verbally and via a 5-point Likert scale indicated that people talking to the CA framework enjoyed the conversation, found that the CA answered mostly appropriately, and that the conversation prompted memories of past times. Competency questions were also used to check that the reminiscence ontology answered correctly.

Spell-Checker

The function of this was to make an attempt at understanding what a person has said to the CA even if they had spelt a word incorrectly. For example, 'what will the wether be like today?' would not be interpreted as weather without the spell-checker facility. The spell-checker is designed so that words that are spelt incorrectly can be checked against the spellcheck file and then correct spelling substituted within the CA system and the correct subject matter is then returned to the person conversing with the CA. If a used word is not recognised by the system it is added to the unknown words file for adding and amending to the system manually. The spell checker is populated with a glossary of things from the past,

such as ‘chitterlings’, ‘clackers’, ‘ha’penny’. The spell-checker comprises many variations of word spellings with wildcards used to cover other variations. People miss-type and can make spelling mistakes that would prevent a CA from responding correctly. The spell-checker feature provides a way of providing a correct response or reverting to a holding reply if the content cannot be interpreted.

Short-term memory function

Recalling the last 25 iterations to allow some interactive unscripted interaction with a person talking with the CA. Also, recalling the name of the person conversing with the system and therefore appear more friendly to the participant. The saving of input is limited to 255 characters per 25 iterations, there is potential to increase this, however, testing of the CA has found this to be currently unnecessary. It is saved in a re-writeable text file that is progressively overwritten each time a new iteration occurs, totally refreshing every 25 times. The conversation is carried out between a user and the CA. The CA knows the name of the person through their logging into the system. The CA is either set to short-term memory only mode or ‘Learn’ mode. Learn mode is the long-term memory functionality. In short-term memory mode only the last 25 iterations are reused in the conversation with the CA and once used are discarded for the next iteration. The short-term memory function passes unknown keywords to the system so that they can be utilised by the system in the ontology of reminiscence later. Having the short-term memory mode on without the ‘Learn’ mode results in a more static conversation.

Long-term memory function

This functions by remembering past conversations so that people can actively participate in interactive conversation. A database of past conversations that can be accessed by an individual during further conversations with the CA. Shared reminiscing of some content such as ‘cleaning the step with a donkey stone’. So that when donkey stone is

added to the reminiscence ontology it will available to all participants should they refer to such a thing. A time limiter function was built in so that during experiments a set time for interactive conversation was set. Reverting to the short-term memory mode thereafter. This functionality was successfully implemented and accessed by the correct logged in individual. Over time creating a personal story of a person's past as discussed with the CA framework. The CA functioned well and proved to be a robust framework. Conversational logs indicated successful interaction between people and the CA. Feedback given verbally was for great enjoyment and satisfaction with the system as a virtual companion to some of the participants.

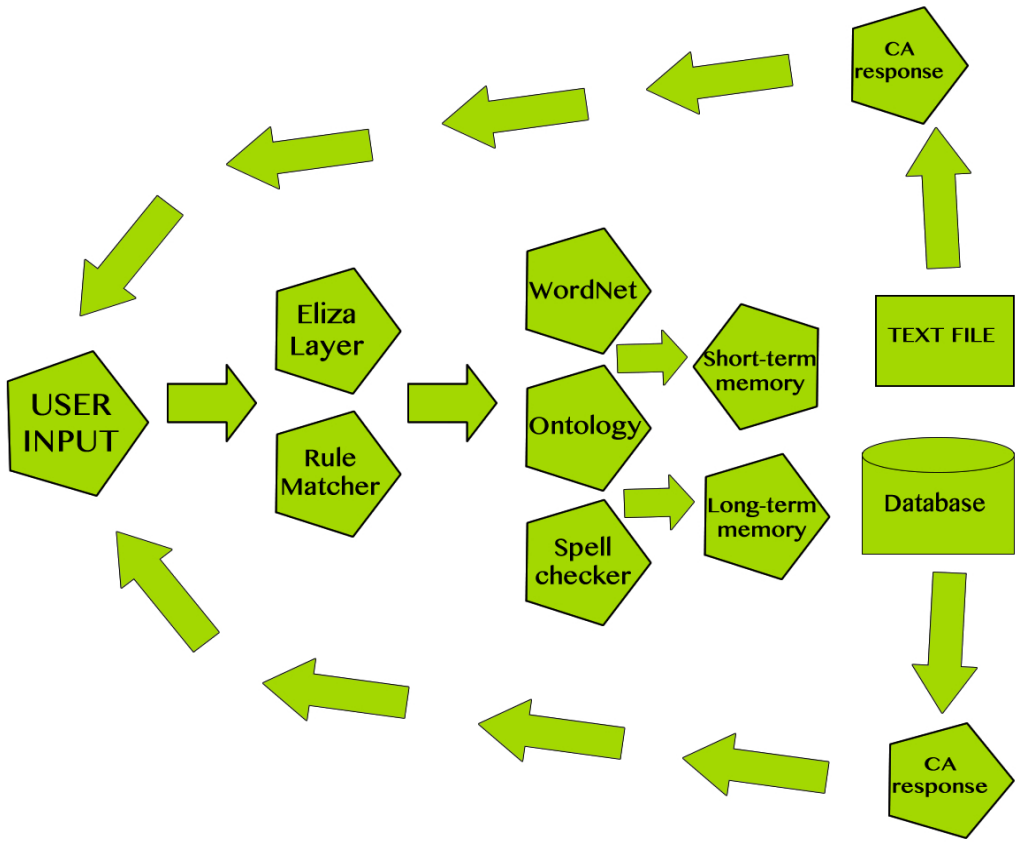


Figure 4.1 The reminiscence conversational agent framework architecture design

4.3 Reminiscence CA framework design

The architecture of the baseline reminiscence CA framework is illustrated in (Figure 4.1). After appraising other CA systems, including InfoChat, it became clear that any CA system would require certain standard components. These could include a form of text or speech recognition capability, natural language understanding to analyse input and produce a representation understood by the system, followed by a dialogue manager to keep track of information received, then a task manager to process the information and return the output via a natural language generator to text or speech output. Due to the CA system being designed for use by older people to reminisce, a themed domain was necessary and that is where the idea of creating a reminiscence ontology as well as scripted historical themed stories came about. It was also necessary to consider use of such a system by older people in terms of usability and accessibility. It was also important that the system didn't appear off putting to older people with sometimes low levels of computer skills or confidence at using such systems.

The baseline system follows the following sequence of events:

User visits web page or logs onto CA locally

Log in results in tailored content for the user. The prototype CA framework is generic until log in by authorised users.

Example Web page sequence:

Agent greets user generically	*Hello, how are you today?*
User types in answer	*Hi, I am very well today.*
Agent responds to answer	*I am glad that you are well.*

	What shall we talk about today?
User types in answer	*I'd like to talk about my schooldays.*
The agent looks for a theme based on what the user has inputted.	*Did you like school?*

What happens during the conversation?

The CA needs to parse natural language to understand what is being said or queried. The system needs to figure out if what is said is a positive or negative fact or opinion. It also needs to remember what has been mentioned or asked to reply with a coherent sentence. When the CA has a question asked, it needs to consider its conversational history or other knowledge base and extract an answer that is relevant and generate a statement that grammatically works and doesn't contradict earlier statements. If a preceding step fails, then the CA needs to ask a clarifying question. It needs to keep track of the conversation structure so that incorrect comments are not triggered. The pre-fabricated general remarks need to be used to maintain the communication and encourage the user to say interesting things.

Conversation manager

- The conversation manager takes the user input and passes it through the Eliza layer and Rule matching layer to the novel Reminiscence ontology. The system automatically performs

spell-correction, idiom substitution, proper name and number joining. It strips off trailing punctuation but remembers whether or not the sentence had an exclamation or a question mark. The control script will later examine the construction of the sentence, so questions lacking a question mark will still get appropriately marked.

- The CA consists of rules organised into contexts and topics. Some rules are used to match arbitrary user input, like *user: where do you live?* Others are used to lead the conversation on a subject, like *Betty: do you have any pets?* And yet more rules are used to react correctly to a user's response to the CA, like *CA: What is your favourite way to cook eggs? User: I think scrambled eggs are best. CA: Isn't it better to eat boiled eggs?*

- Topics typically are a collection of rules based around a conversational theme, be it food, or pets, schooldays, the Sixties or family. The CA system is very flexible in that you can write a topic about anything and implement it into the framework system and have it automatically linked in and working. So potentially lots of people could collaborate by picking a different topic to write about and compiling them all collectively into the system.

- Better topics will assist greater interaction with the user and better share personal opinions and information. Interacting with the user will mean that the CA asks the user questions and volunteer our answers to those same questions. If the CA asked something like "*do you like eggs,"* it is expected the user will give their opinion on the subject as well – a sharing of information.

Whether the CA tells the user one view first and then solicits theirs or asks them and then volunteers the CAs is a matter of choice. For more natural interaction the CA will vary its style so that the order of questions changes at times. But it appears better to ask the user first and answer with the CA second. If you tell the user a fact first, it often stifles their opinions and leads to shorter interaction with the system. If the user has knowledge of

what the CA is talking about before asking a question, they may volunteer their answer before the CA asks the question. Such a system becomes a tricky scripting problem because the system would have to detect they did this (because it would be pointless to ask the user after they have just told the CA).

- One of the good things about asking questions is that the user response is to answer directly. Making it simpler to write rejoinders that can interact with them. *"Do you like swimming?"* Obviously asking yes-no questions makes it easier to write responses. But too much of that will cause the conversation to appear more artificial. Conversational variety offers choice questions like " *Would you rather eat a scrambled egg or a fried egg?"* Though still limited to two rejoinders. The really good choice questions appear to leave the options to the user, like the earlier "*What's your favourite way to cook eggs?"* There are lots of ways, and the range of choices are covered in about 20-30 rejoinders. Having a good response to what the user says will make the CA seem more intelligent and responsive to the user.

Response generation

This element produces the response after the following layers have activated

Eliza layer

- The Eliza layer ensures there is always a response from the CA and an attempt to steer the conversation along with a planned path of replies.

Rule matching

- If elements match rules, the responses are formulated, and certain paths are followed within the conversation.

- For each word (regular and canonical), the system chases up the fact hierarchy of them as input patterns and the possible variations allow the system to continue processing. As it comes to a word or a replacement word, it marks on the dictionary the position in the sentence where the original word appeared. It does this same thing for sequences of contiguous words found in the dictionary. User input is matched and an appropriate response given.
- Rules are organised into contexts as well as topics. Topics provide themed batches of conversational material. Contexts provide a heading for the system and allow the system to categorise the input and return an appropriate response.

Reminiscence Ontology

- The knowledge base for initial themed conversation and subsequent threads to the conversation when it is activated.

- Allows shortened lines of code in the script by using the ontology as a dictionary of terms and then referencing back to them in the script.

- Typical user interaction is the following: (i) the user enters a query, e.g., what's a donkey stone? (ii) system navigates within the reminiscence ontology to refine the query, (iii) system abstracts concept from ontology, (iv) system returns definition in appropriate format. e.g., It's a handheld stone for polishing a doorstep, popular during the 1940's – 1960's.

- Data within the reminiscence ontology is contained within a hierarchy of tagged classes which are also mapped to WordNet. Associations between these tagged classes are automatically acted upon by the system. They are also added to as new data is brought into the system. Currently this is done manually, but in future will be automated. Defining the ontology of reminiscence uses a set of known classes which are subsets of the reminiscence domain. The

reminiscence domain subsets are split into attributes and roles. These subsets have a number of assertions which restrict the number of possible meanings within the reminiscence ontology.

- The ontology search algorithm takes the user input, searches for named classes that match the keywords detected, then it looks for related terms which refine the user input. Matched input is then returned with the most likely output based on how the input was interpreted to appropriately respond to the user input.

WordNet Ontology

- The WordNet ontology is invoked by naming the word and meaning required.
  Concept: ~buildings (shelter~1 living_accomodations~1 building~3)
  The concept ~buildings represent 760 general and specific building words found in the WordNet dictionary – any word which is a child of definition 1 of shelter, definition 1 of accommodations, or definition 3 of building in WordNet's ontology.

- Allows more conversational themes to be explored giving a range of possible topics with which to converse. Starting conversation is often difficult for people so having a ready supply of subjects to talk about should help.

- Provides a lexicon of words and meanings for use with the CA

- Is an ontology in its own right

Database

- Storing of user name and previous conversational matter will enhance the overall experience for the individual participant

- The knowledge base increases in size the more people converse with the CA. The knowledge base can be switched off for general

lower level conversation and turned on using 'learn' mode for more in- depth and long-term memory capability of the CA. There is currently no actual size restriction placed on the growth of the system, however, depending on where the system is running, and on what interface, this could be a concern due to lack of available memory on the user's computer or the web server if a web interface.

- Short term memory - User details

- Storing user details will help to maintain current conversations and add to the overall conversational experience of the user

Long term memory past conversations

- Storing and accessing previous conversations with a participant will enhance the conversation

Spell checker

- Helps identify words within the knowledge base and ontologies and assigns the correct spelling of the words so that the correct meaning is attached to the word(s).

Verb ontology

- Provides a source of verbs for the CA system that was not present in the WordNet ontology.

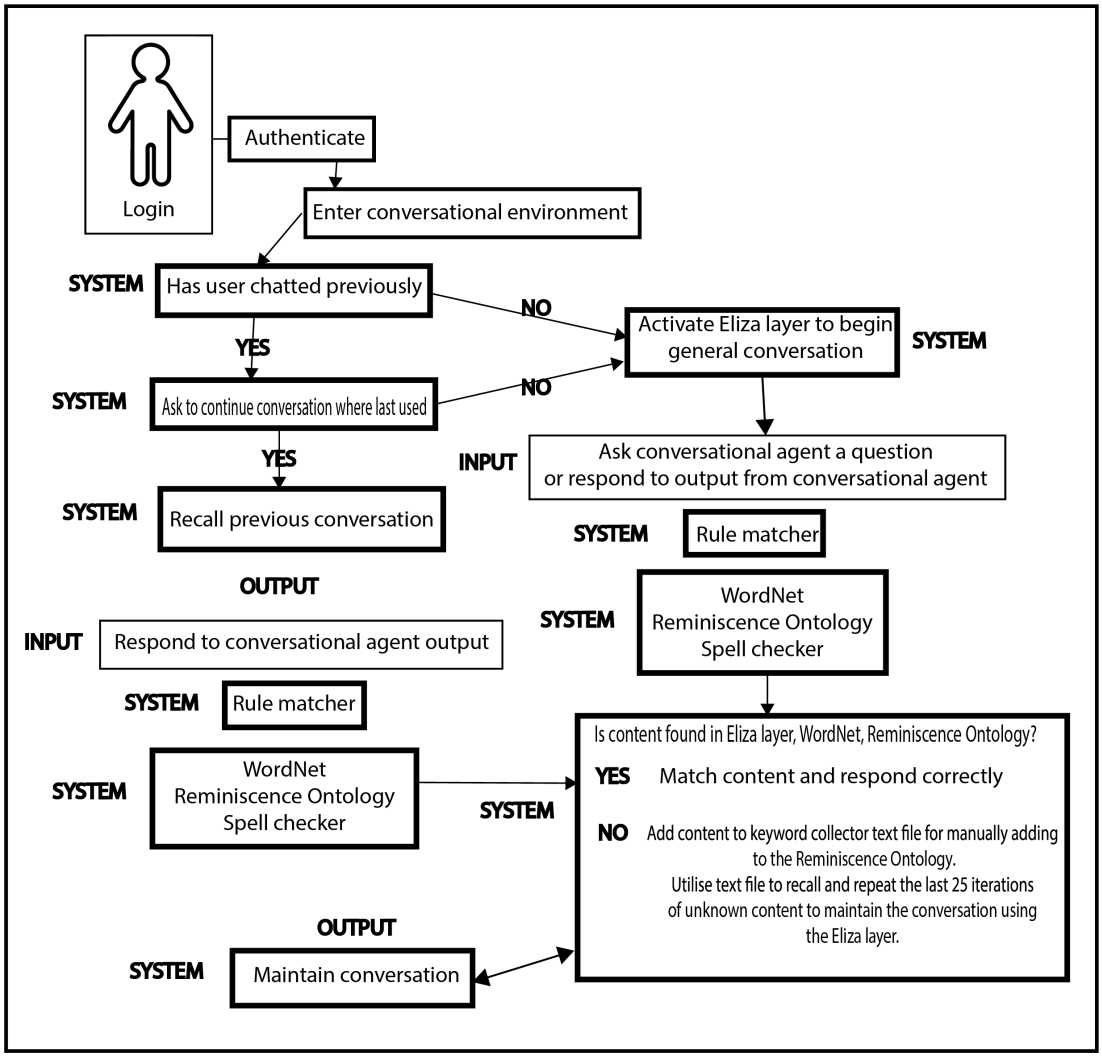


Figure 4.2 Generic diagram of CA system showing how the system works. The user logs in and is identified by the system as an authorised user. The system checks whether there have been previous conversations. System maintains the conversation.

4.4 CA System Design and Implementation

A generic diagram of the system is illustrated in (Figure 4.2). This shows the user logging in and being recognised by the system as an authorised user. If they haven't spoken to the CA previously, the system activates the Eliza layer to begin conversation and to illicit responses from the user These responses are passed through the Eliza layer and rules activate which fetch an appropriate response from the system. If there have been previous conversations with the system, the last topic is recalled by the system and the user is invited to continue where they last spoke. If the person wishes to revisit past conversations, then they respond positively

and the CA system will proceed to use the correct conversational material saved in long-term memory. A negative response to continuing the previous conversation will result in the Eliza layer activating to promote conversation and elicit further conversation with the person. All spoken content is passed through the Reminiscence ontology, WordNet and checked by the spell-checker. The short-term memory of the CA recalls the previous 25 iterations to enhance the conversation. The long-term memory when set to 'Learn' mode saves all conversation to the knowledgebase after extracting keywords into a text file of words that are not recognised by the system. This allows new words to be manually added to the Reminiscence ontology for future use during conversation. This could be automated as long as there is a mechanism for checking that unwanted conversational topics such as abusive conversation isn't added to the system. Currently, all new topics are added manually. Some of the new topics gathered include local names and terms for things, such as, 'moaning minnie'. This was the nickname for air raid sirens during World War Two and is generally used to describe someone who keeps whinging. Other subject matter included, 'donkey stones', 'chitterlings' and hap'orth o chips'. It is interesting to note that the system saved a lot of northern phrases and terms which enhanced conversation for the local people taking part in the studies undertaken. It may be positive to have regional ontologies for different areas, so that the conversation content is relevant to each person taking part in conversation with the CA system. These would all be triggered by a person saying the relevant word or words during conversation, so no one would be excluded from potential conversational material.

Testing and evaluation of the system was done by checking the conversational logs to see what rules fired and what was said to the CA and the CA given response to the conversation. Also a 5-item Likert scale was utilised during the later experiments to gain subjective insight into the CA behaviour and impressions of people talking to the system. Testing for usability early on in the experiments (Chapter 5) informed the look, behaviour and sound of the CA system.

It is proposed to do further analysis of the conversational logs to gain metrics of success rates of conversation carried out by the CA. This would be done in future work and is suggested in chapter 7.

Scripting mechanism

This was developed to access the reminiscence ontology as well as the WordNet system. It was designed so that shorter scripting was able to be used for creating the system. The CA system accepts username at login in a CHATLOG.NET file. This is collected as a value ‘name-0’. From then on this value can be called as required by the system. So for example, the CA system might firstly say ‘Hello! My name’s Betty. Is your name #0?’. (Where #0 was set to users name). The name parameter was collected on login. The person can correct their name and choose a nickname or they can answer ‘yes’. Successive returns to the CHATLOG.NET file during conversation elicits different responses to be used during the conversation. These are shown in init-chatlog-*. (Where * is the number in response playback order, 1-999 etc). The greeting will change according to time of day, number of times logged in or simply use an alternative to the ‘Hello! My name’s Betty….’ iteration. Random responses are included here to vary the introduction to the conversation. A person returning to the conversation will never encounter the CA informing them of the CA’s name as they are assumed to already know this from previous conversations. The person can ask ‘what is your name?’ and the CA will then offer that information.

CHATLOG.TXT is a log file of the latest conversation showing the rules that fired, so this can be checked, and rules amended as necessary to enable more coherent and seamless conversation with the CA.

During testing, TRACE can be active to show on the screen what rules are fired and their order. This is useful to test the CA is working correctly. These messages are also saved in the log file for later analysis.

The Quit String of the system is set to ‘Bye’. Because the system is mapped to WordNet, any permutation of ‘Bye’ given will end the current conversation and log the user off. The system qualifies that the user

meant ‘goodbye’ before actually logging them off. There is also a time delay for inactivity that causes the CA to ask if the user wishes to talk about something else, continue the conversation, or leave the conversation. A null response will result in a log out happening after 1 minute of inactivity. This null response rate can be lengthened if required but is set at 1 minute for the experiments.

Rules firing

In the Eliza layer of the Reminiscence CA, all rules can be re-fired immediately, in the other layers, rules are set to only re-fire after a time delay so maintaining a fresh conversation and not seeming to repeat general content within the same conversation. People rarely use grammatically correct or complete sentences in conversation. Therefore rules manage the context of the conversation. When new rules are added they have to be checked to ensure they interact with existing rules correctly. So rules within the Reminiscence CA are segregated into ‘contexts’ that limit the rule scope and result in better parsing of the conversation. Thereby producing an appropriate response to the user input. For example, if the user includes the word ‘Football’ in their sentence. Then the context of the conversation is ‘Sport’ which makes the possible responses limited to those about sport, rather than about other childhood games. However, being mapped to WordNet and the Reminiscence ontology, the conversation will proceed as further rules dig deeper and activate in this ‘football’ instance, to more possible games or activities played, matches visited or played in and history of football clubs like Manchester City and Manchester United. A conversation based around the keyword ‘football’ could develop into reminiscing about the ‘Busby Babes’ or England’s World Cup win in 1966.

The rule groups are placed into contexts, of which one will always be current to maintain conversation. A user input is matched with the pattern in each rule, pattern matching. The initial match fires the rule’s response and the corresponding rule is invoked. Filters deal with changing conversational topics. The filter has keywords that indicate that

the user has changed topic. Further rules activate to attempt to elicit further input that can be recognised by the CA system and respond appropriately to the new conversational topic.

Storytelling potential

The Reminiscence CA has storytelling potential. Set storylines are included within the system that lead the user in conversation about the current historical theme. Themes currently available include 'The Sixties', 'The War Years' and 'Schooldays'. These themes break down into more categories and are activated by the user's interaction with the CA. For example, 'The Sixties' can ask if the user liked music in the sixties and who they liked. Bands and songs as well as the Number 1 hits for the whole of the Sixties is available to enhance the conversation. Cars of the period, such as the iconic Morris Mini Minor and Monte Carlo Minis are available to talk about. It is evidenced from the conversational logs that people talking with the CA framework enjoyed these themes and took part in conversation about them. More content can be gathered from the logs to increase the conversational content of the Reminiscence CA as more people engage with the system. The rationale of using this approach is that people like to reminisce and is something we all do (Gillies and James, 1994).

As related in chapter 2, reminiscence is a positive experience that can have a direct influence on subjective wellbeing and also memory recall to aid improvement of normal age related memory problems. Reminiscence encourages conversation and is seen as a fun thing to do (Bryant et al., 2005).

4.5 Performance of the algorithms

The overall performance of the CA was very good. Response rate per CA iteration was no more than 5 seconds. So people talking with the CA did not experience great delays in gaining meaningful responses.

The system has a delay before running when it has been re-compiled due to new data or program modifications being made. These were always

undertaken during times when people were not expected to be logged on. This was successfully done, utilisation of a testing version of the CA and also the live version meant that re-compilation errors were controlled.

The rule matcher was modified repeatedly until it reached a satisfactory level of quality. Some rules were never fired and this could mean that some rules were already redundant when the CA was live. Subject matter will vary from 'place to place' as reminiscence experience and people's life history varies and so redundant rules may be activated by different people from different areas conversing with the CA system. The collection of unknown data would benefit from being an automated system but there would still need to be a checking mechanism to ensure inappropriate content isn't added to the CA knowledgebase. Saving of content and later sharing some of it with others via the Reminiscence ontology proved invaluable and provided further content for interactive discussion. The Reminiscence ontology was produced with a set of terms useful for people to reminisce about. This system was expanded as new information was collected. The data collected was mainly related to Northern England as all the participants were from Northern England. Local names for places and items remembered proved topical for the group of people testing the system but would be different for people from different regional areas. Therefore the Reminiscence ontology production technique would need to be carried out in each new region to collect initial data for the ontology. WordNet requires a supplementary system to fit within the reminiscence domain. It is in itself invaluable as a tool and it is thanks to Princeton University that such a tool has been made available. The spell-checker was also necessary, older people do not always type accurately and are not always over-confident using computer systems. The education level of the participants varied and some had left school at 15 with no formal educational qualifications. They did however have a wealth of life experience that could be utilised by the CA system. Had the system not responded correctly to a spelling error or typographical error, then the participant would have had a negative experience of using the CA system and may have been put off returning to

use the system. The spell-checker did not correct their spelling but rather answered them correctly and maintained the conversation on the theme. The conversational logs provided a quantitative measure of the overall performance of the CA system both in terms of time taken to respond to input and in terms of accuracy of that response to user input. Allowing adjustments to be made to the CA as necessary to improve its performance. The logs also indicated which topics were more popular and therefore informed the design of future topics for conversation.

Summary

- The CA system supports a simple visual syntax and powerful pattern-matching features. Matching concepts instead of single words and being able to specify words that must not be in the input allow the CA to approximate patterns of meaning. Using ranged wildcards limits false positives. Subdividing rules into topics accessible via keywords makes large collections of rules efficient to search and makes it easy to author orthogonal content.

4.6 Defining the reminiscence ontology

Possible interpretations of the term “ontology” (Guarino and Giaretta, 1995) are illustrated (Figure 4.3).

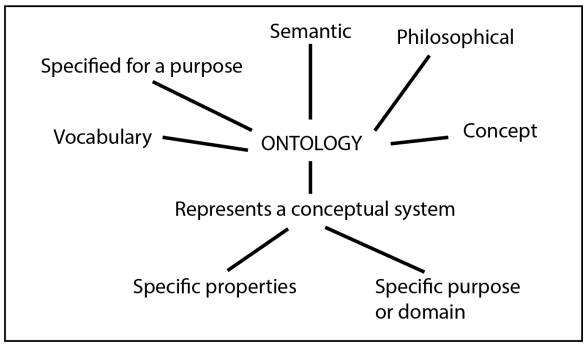


Figure 4.3 Ontology systems comprise many components. Representing a conceptual system and thereby a specific domain ontology will create a Reminiscence ontology.

Quickly understood by the research community is the necessity of creating specific purpose or domain ontologies in a semi-automatic way (Navigli et al., 2004). Using domain ontologies for knowledge representation has two main advantages: firstly, their standard formalism enables sharing and reusing ontologies between any ontology-friendly environment. Secondly, their formal structure allows the possibility to reason over the obtained knowledge representations and to envisage the automatic extraction of the ontological components as modular layers. This automatic ontological extraction is known as "Ontology Learning" (Navigli et al., 2004). Domain ontology defines relationship among classes, properties and instances used in the themed ontology. Once the concept of ontology learning was defined, production of the Ontology of Reminiscence began. Initially, the text was written down in natural language that described the reminiscence domain. The natural language text produced from data collected during face-to-face group discussions with the pilot group. The collected data allowed the creation of a glossary of natural language terms and definitions (Butler, 1963). A preliminary ontology was created and mapped to the WordNet hierarchy (Miller, 1995), (Fellbaum, 1998), and then implemented within the CA. By breaking down the ontology into nouns, adjectives, opposites, prepositions, verbs, and concepts, the mapping to the WordNet hierarchy followed. The identification of concepts was the first task to be completed. Concepts in this instance are complex mental objects that are characterized by numerous features. Concept extraction refers to the identification of important domain classes.

In the terminological approaches, concepts are terms that are particularly important for the domain. These terms are extracted from the corpus as outlined by (Buitelaar et al., 2006) who consider that a concept should have a linguistic realization. In this case, the major challenge is to be able to differentiate domain terms from non-domain terms, usually using statistical filtering. The identified terms (composed from single or several words) are considered as concepts/classes, or

they are classified according to broad classes already available in thesauri and vocabularies. Other approaches rely on clustering and machine learning to learn semantic classes. In this case, a concept may have no corresponding term in the corpus.

An ontology life cycle example is schematized by four main stages: the specification stage, the formalization stage, the maintenance stage, and finally the evaluation stage (Price and Spackman, 2000).

- The specification stage allows the identification of the purpose and the scope of the ontology. The specification stage is primarily based on domain experts and requires the definition of competency questions that the ontology must be able to answer. It is also dependent on the application that is going to use the ontology

- The formalization stage produces a conceptual and formal model that satisfies the specification stage

- The maintenance stage allows ontology update and evolution and checks its consistency
- The evaluation stage analyses the resulting ontology and checks if it meets the initial needs and if it has the desired features.

With the reminiscence domain ontology, it was of particular interest to focus on the formalization stage and in how this stage can benefit from automated methods for knowledge acquisition. Hence, it is interesting to state explicitly the automated steps that alleviate the task of human experts and the burden of knowledge acquisition. Ontology learning techniques can be adopted to reach this goal (Aussenac-Gilles et al., 2000). These learning techniques can vary according to the degree of automation (semi-automatic, fully automatic). Also, the ontological knowledge extracted (concepts, taxonomy, conceptual relationships, attributes, instances, axioms), the knowledge sources (texts, databases,

XML documents, etc.). Finally the purpose (creating ontologies from scratch and updating existing ontologies).

Defining an ontology of reminiscence - this followed a suggested ontology production technique (Fellbaum, 1998).
This production technique was followed, then modified as illustrated and an ontology of reminiscence was created, see Table 4.1

Table 4-1 Ontology production technique

Make a list of relevant keywords gathered from conversation with participants in face- to-face groups [Appendix 13]
Collect keywords from conversational logs of interaction with the prototype CA [Appendix 14]
Relate keywords in a spider chart (Figure 4.4)
Begin entering keywords [Appendix 1]
Link keywords and attributes
Plot the *Has a*, *Is a* type relationships
Manually prepare the data into a format readable by the CA. Create a semantic triples environment. A triple set of three entities that codify a statement about semantic data in the form of subject-predicate-object expressions. For example: “Collette is 42” or “Collette knows Jim”. The subject is the object the triple is describing. Predicate defines the piece of data in the object we are giving a value to and the object is the actual value. This technique enabled collected data to be represented in a machine-readable format. The coded system also included using stored arrays which

are not currently by other semantic web systems. The system dynamically creates and stores triples during active conversation.

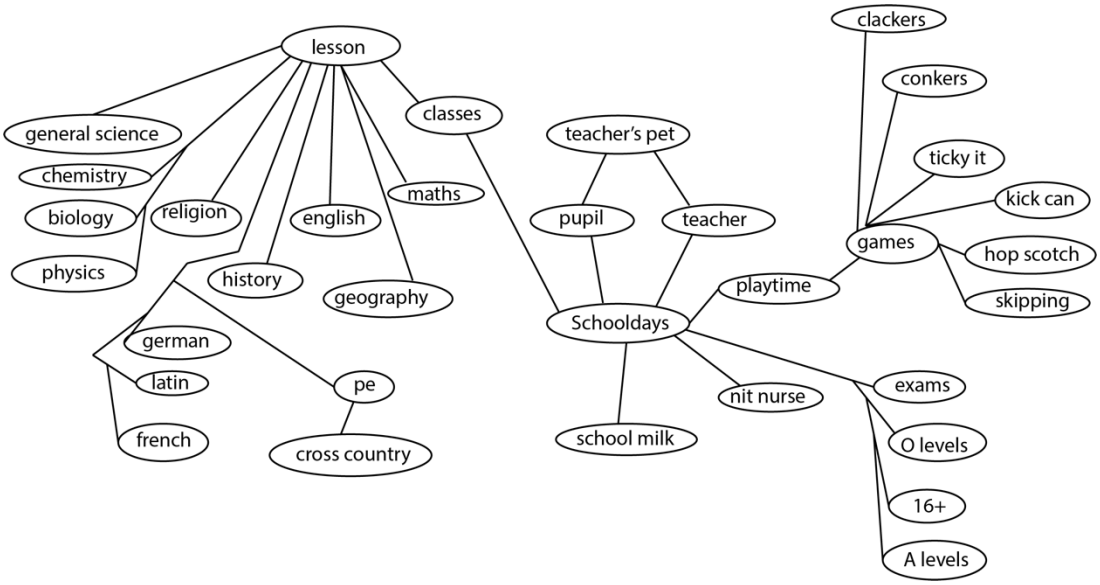


Figure 4.4 Ontology hierarchy demonstrated for several branches from Schooldays.

The reminiscence ontology was then applied to the Reminiscence CA framework, See Table 4.2 below.

Applying the ontology –

Table 4-2 Applying the reminiscence ontology

Import the reminiscence ontology into the CA memory. Importation required breaking down the ontology into nouns, adjectives, verbs and concepts.
Test the robustness of the ontology framework during conversation between the CA and the participants via the conversational log files. Test the ontology by asking a list of 20 competency questions (Table 4.3). Content was held within the Reminiscence ontology about World War II and other themed subjects. Competency questions are defined to evaluate the quality of a given ontology. To be of value all

questions must be answered correctly.

Table 4-3 Competency questions to test the reminiscence ontology.

Competency questions	
What was a doodlebug?	Who won the Monte Carlo Rally in 1964?
When did pupils leave school in 50's?	Who was King during WWII?
What did 'Make Do & Mend' mean?	When were children evacuated?
Who was given free milk?	Where were children sent to in WWII?
What is tripe?	What was a ha'penny?
Who were the LDV?	When did Winston Churchill die?
When was the first Moon landing?	What year did sweet rationing end?
What is a coal hole?	Who were the Beatles?
What is the name for an air raid siren?	What was the Land Army?
What was made at Avro's?	What was the blackout for?

Evaluating the ontology -

Evaluate the ontology versus a hand-crafted method for ease of use, the speed of production, utilisation of data from conversational records. Hand-scripting each possible permutation takes considerable amounts of time in preparation (Table 4.4). Use of the ontology cuts out some of this time and makes the conversational script easier to follow.

Table 4-4 Evaluating the ontology

The text was written down in natural language that described the reminiscence domain. Natural language text collected from projective interviews with participants before using the CA. It allowed the creation of a glossary of natural language terms and definitions (Curry et al., 2011).

During the pilot phase, three distinct but intertwined ontologies were

identified and in two cases created - reminiscence, verb ontology, and the WordNet ontology.

The reminiscence ontology was created using keywords initially, these were gathered from questionnaires and focus group discussions. Here are two examples of material gathered from the reminiscence questionnaire [Appendix 4].

"I remember the 60's. It was a time of brilliant music and great bands like the Beatles were around. I saw them play at the Cavern in Liverpool. I lived in Bradford, Manchester at the time on Keywest street. The street was pulled down during the early 70's. I remember the money when it changed to decimal in 1971. A shilling was worth more than 5p so we lost out. When I was young I was given a penny a week as spends. Imagine giving a kid today 1p a week".

"I remember the blitz as though it was yesterday. We had no street lights at night so when the air raid siren went off we had to go outside in the dark. It was cold and damp in the shelter I used to put my hands out in front to feel the way. The siren was called a moaning Minnie and that's what my mam called us kids when we were moaning on".

From this material, entities were identified to go into the reminiscence ontology. Relations to the entities were mapped out and then specified relations between the entities were established (Figure 4.5).

Figure 4.5 Itentified entities were added to the ontology framework and relationships attached to allow the ontology to utilise the data when requested.

Meaningful relationships between entities were established and using the scripting mechanism described earlier, appropriate conversation between a person and the CA system took place.

4.7 Mapping of the ontology of reminiscence to WordNet

WordNet consists of sets of orthographic forms of words with associated sense sets: a set of synonyms or synset, which define a concept. WordNet also sets out a large number of domain-independent lexical relations which can hold between synsets, e.g. hypernym (is-a), ‘part-of,' ‘member-of.' These connections can be linked together to form paths (e.g. hyperonymy chains) between synsets (Milward and Beveridge, 2003). The CA knowledge base derived primarily from WordNet, which has some 150,000 words, creating a hierarchy (ontology) of words. The system currently has a complexity limit set to allow for growth of the ontology during use and can handle sentences of 255 or more words. This is more than usual CA user utterances, so should not cause a problem (Milward and Beveridge, 2003). Extensive lexical ontologies such as WordNet have been used in applications such as text summarization (Chaves, 2001). They have been used in

translation (Knight, 1993), and also, information extraction (Burke et al., 1995), sometimes directly, sometimes to help build more specialized domain ontologies (Milward and Beveridge, 2003) and for several other diverse applications. Ontologies have been used within specific dialogue system modules, for example in language generation as the basis of the Systemic Grammar approach (Bateman, 1990). However, more general use of ontologies is relatively rare (Milward and Beveridge, 2003), (Kluwer, 2011). Words were marked with concepts and then linked to the WordNet dictionary. Some existing terms were expanded [Figure 4.6 & Figure 4.7] .

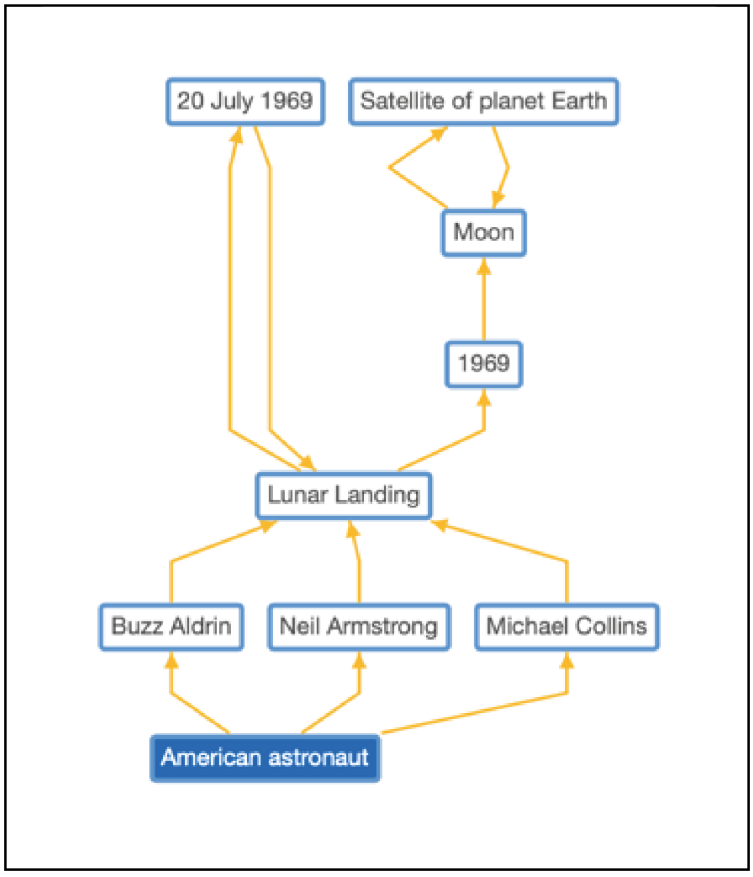


Figure 4.6 In this entity graph, items relating to the lunar landing of 1969 are shown. WordNet had Moon, date of landing, Neil Armstrong and American astronaut listed, other items were mapped to this through the reminiscence ontology.

```
table: ~memory (^role ^memory ^kind ^item)
^createfact (^memory member ^role)
if (^kind != *) {^createfact (^item member ^kind)}
if (^item != *) {^createfact (^item write ^memory)}
^addproperty (^memory NOUN_MEMORY ) // Map to WordNet
DATA:
~news Neil Armstrong ~event [Lunar_Landing_1969] //Map to WordNet
~news "Buzz Aldrin" * *
~news "Michael Collins" ~event "First moon landing"
```

Figure 4.7 Mapping to WordNet, Moon landing 1969

Parsing was done using a pseudo-parser, Part Of Speech (POS) Tagger, which attempted to label how words were used (Kluwer, 2011), this seemed more efficient than using a pattern matching algorithm and will be tested further after the pilot phase has ended.

A novel contribution to the development of the CA is the reminiscence ontology and the personal user ontology that supplement WordNet to expand the available ontology. The bespoke verb ontology and the integrated spell-checking facility enhance the capabilities for the user. During a conversation with participants, each term in the ontology is activated, and by studying the resultant conversational log files of the conversations, the ontology was further refined and consolidated.

The reminiscence ontology was built by collecting keywords from oral discussions with a group of people aged 60+ with NM, a written three question questionnaire and also from conversational logs of interactions with the pilot CA framework. A mapping approach was followed to align the ontology with the WordNet system (Figure 4.8). Where content didn't exist within WordNet, this was added to the reminiscence ontology.

```
evacuation ( meanings=3 glosses=2 NOUN_ABSTRACT  NOUN  NOUN_SINGULAR  COMMON2  COMMON1
NOUN_NODETERMINER posdefault:NOUN )

        expunging~1n
        evacuation~2nz  (^safe~2) sent away to a safer place during World War II
        evacuation~3nz  (^rescue~5) to remove something quickly in an emergency
```

Figure 4.8 Mapped to WordNet format and saved in a form readable by the CA system.

Hierarchical relationships between concepts and instances were considered. The ontology is a hierarchical representation of classes, sub classes, their properties and the instances. The reminiscence ontology was created as a domain ontology. It was inspired by work on the semantic web using RDF and OWL.

An example of a dialogue with the CA, illustrating some of the concepts utilised, is given in the following table (Table 4.5).

Table 4-5 Example dialogue with the CA

CA	**USER**
Hi there what are you called?	I am Collette.
Collette, pleased to meet you	And you too.
What would you like to talk about?	I don't know.
Tell me a bit about yourself.	

The CA memory allows the system to identify when a user has returned to a conversation after a period as shown below (Table 4.6).

Table 4-6 The conversation when a user returns to the conversation after a delay

CA	USER
Hello again, welcome back.	Hi there.
Shall we continue with our last conversation?	What was that about?
We talked about your rabbit, Nibbler.	No, let's talk about something else.

During the dialogue, the user may provide a response to a prompt that does not match any of the expected range of replies (as defined by the reminiscence domain) but is a hypernym or hyponym of an expected response. In this case, the CA should be able to discover the relation between the expected terms and the user response to resolve the discrepancy. In the case where the user's reply was a hypernym of an expected term, then their reply can be considered under-specified, and the CA can issue clarification questions to obtain a more detailed response (Table 4.7).

Consider the following example:

Table 4-7 The CA asks further questions to clarify facts

CA	USER
What relatives do you have? ***The system expects uncles, aunts, brothers, sisters, etc.***	Yes, lots
What type of relative?	Two brothers

In this example, the CA expected a specific answer, but the user replied with a more generic term. The system, therefore, formulated a more accurate question to elicit an answer at the expected level of the is-a hierarchy. Hypernyms are similarly treated in a generic fashion in the system via the use of a disambiguation operation.

In the case where the user's reply was a hyponym of an expected term then, their response can be considered over-specified, and the system can: (a) find a more general related term which matches the expected responses to answer the current question. (b) avoid asking subsequent more specific questions that have been answered already (Table 4.8).

Table 4-8 The CA can answer unexpected responses

CA	USER
What children do you have? ***The system expects son, daughter, etc.***	Yes, two.
Thinks - 'How many children?'	
Asks **- What are your two children called?**	James and Amy.
So you have a son and a daughter. **A millionaire's choice!**	

In this example, the system expected a more definite answer to the question (such as son), but the user replied with a more generic term.

The system, however, could match the answer to the question on the basis that two is a son or daughter, which is in a family. The system avoids asking how many children since this has already been answered and therefore would be pragmatically ill-formed (indicated by a thinks). Instead, the system continues with the next item in its plan. During dialogue the user may provide a response to a prompt that does not match any of the expected range of responses (as defined by the reminiscence domain) and furthermore is not a hypernym or hyponym of an expected response, but is associated with one or more of the expected responses by a non-is-a link which imposes an ordering on concepts (e.g. part-whole links). In this case, the system should be able to discover the associative relation between the expected terms and determine whether the supplied term is more or less specific than the expected terms according to the ordering imposed by the relation. In the case where the user's reply can be considered underspecified on the expected terms, the system can issue clarification questions to obtain a more specific reply (Table 4.9).

For example:

Table 4-9 The CA can issue clarification questions to elicit an accurate and sometimes expected response

CA	USER
What type of school did you go to? ***The system expects primary, secondary, etc.***	In the village
How old were you when you went to school?	I was about 5 when I started there.
So, it was your infant school.	

Was it nice there?	

In this example, the system initially expected a more specific school type than the user supplied but recognized that the supplied term village was related to the expected terms in a part-whole hierarchy and was more general than the expected terms. The system, therefore, did not simply repeat the original question 'what type of school did you go to?', Instead, the system formulated a new question to elicit an answer at the expected level of the part-whole hierarchy. In the case where the user's reply can be considered over-specified

On the expected terms, the system can

(a) find a more general term which matches the expected responses to answer the current question and

(b) avoid asking subsequent more specific questions that have already been answered.

For example:

Table 4-10 The CA expects certain answers to questions and anticipates the next question

CA	USER
What type of school was it?' ***The system expects primary, secondary, high, etc.***	In my village
Thinks – 'What was your primary school called?'	

In system-initiated dialogue, questions are grouped together at design time. Even in systems allowing some mixed initiative, this rarely affects the ordering of subsequent questions. Consider, the following dialogue

Table 4-11 The CA looks for responses to queries and offers accurate responses

CA	USER
What is your name?	Collette aged 42
Thinks - What is your age?	
So, you are 42. That is the meaning of life, the universe, and everything!	

Here it is assumed that the system would usually ask for the name first, followed by age. It is intelligent enough not to ask 'what is your age?' once already told, but otherwise, the ordering of the questions is kept fixed. Much more natural dialogues are obtained if questions are clustered dynamically following input by the user.

Consider the following:

Table 4-12 Expected responses to questions result in more natural dialogue

CA	USER
What is your name?	Collette
What is your age?	42

The example shown appears to be a much more natural exchange, with the system immediately asking the follow-on question concerning age. How can this be achieved? Again, ontological information plays a part. Questions which elaborate a previous question, either by asking about a particular attribute or by asking for more specific information (determined by the is-a or part-whole hierarchy) will ideally appear straight afterward, similar to elaboration in well-structured texts (Mann and Thompson, 1988), (Milward and Beveridge, 2003).

4.8 Novel scripting mechanism

To gain the results required of the reminiscence CA it was not possible to use a pure pattern matching approach, the ontology had to be included to ensure that there was a domain-oriented theme and that going off-topic would result in some useful and valid user interaction. The system potentially had a wide range of recognitions of user input alongside expected new terms from reminiscing and misspellings to contend with. This meant the system was to be more complex than a purely pattern matching approach. A purely pattern matching approach would lead to user phrases being unrecognised by the system. The scripting mechanism was designed to take in user input, interpret the input then react appropriately to the user input with no more than a five second delay. The interpretation mechanism works by taking in the input, breaking down the keywords and meanings, passing through the Eliza layer that determines whether the content requires a general response from the Eliza, or if different rules fire, require a specific response to identified content. The passing of data through WordNet, Reminiscence ontology and the spell-checker checks content, returns definitions and triggers the story-telling features of the system. Input content that is not recognised by the system results in the Eliza layer combining with the short-term memory functionality (Figure 4.9). The short-term memory functionality maintains the past 25 iterations.

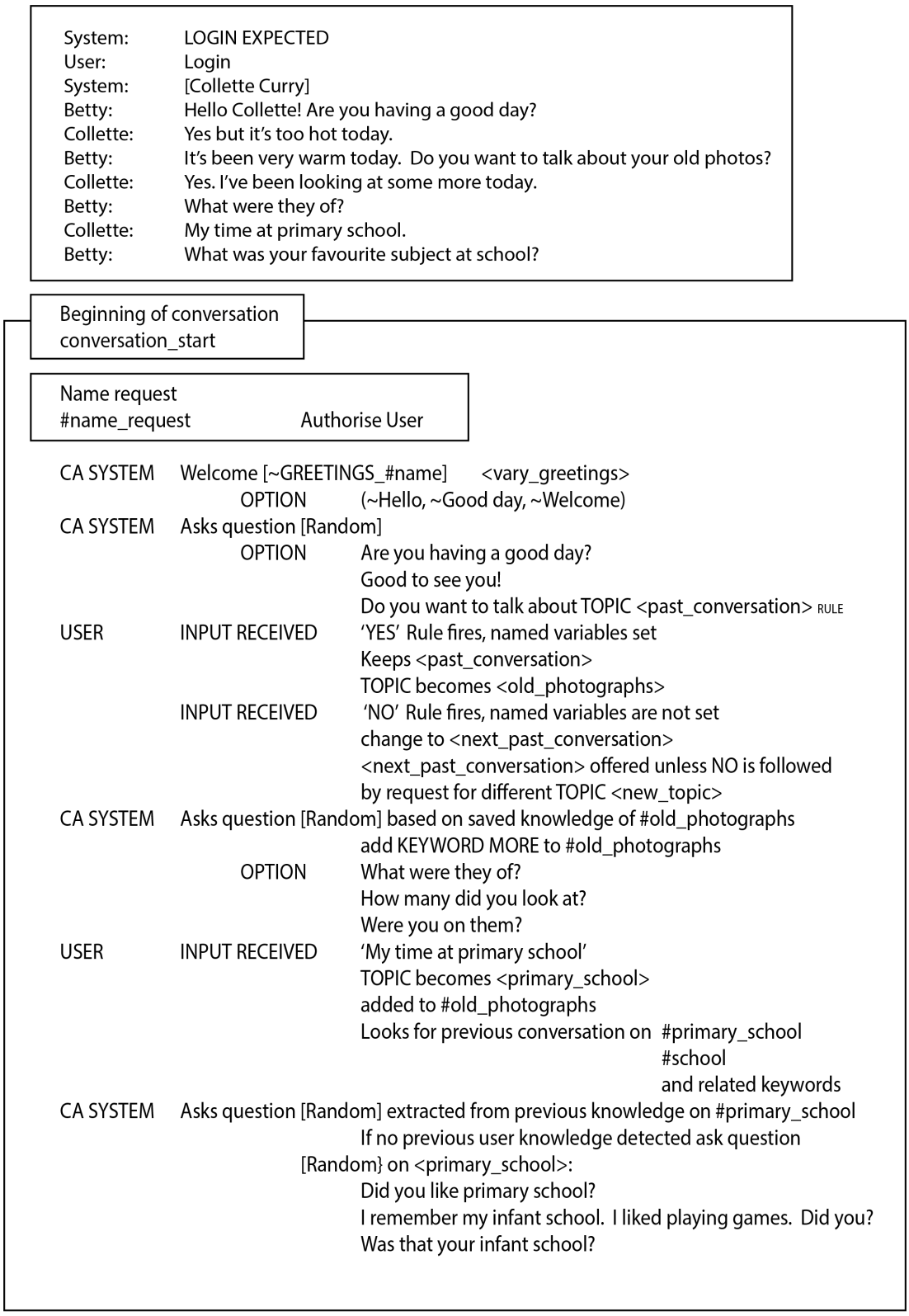


Figure 4.9 The scripting mechanism took in input from the user and as relevant rules fired, content was returned as output to the user.

A novel part of the reminiscence environment is the ability to have a

social conversation with a teachable reminiscence themed CA. A socially oriented chat-like conversation where the reminiscence CA and the user can discuss both domain-oriented topics, such as the Sixties, the War years, school days and working life, and off-domain topics like friends and family (Silvervarg and Jonsson, 2011) is possible.

Reasons for inclusion of such a conversational mode will increase overall engagement and receptivity of the users (Cooper and Baynham, 2005). Also, it will improve recall of the conversational material through emotional engagement (Hamann, 2001), to promote trust and rapport-building (Bickmore, 2003), and to make users feel more at ease with the learning task or topic (Kim et al., 2007).

The InfoChat (Michie and Sammut, 2000) scripting and chat engine programmed in C was initially used to create the CA lower layer with additional functionality and new scripting engine tailor made and programmed in C++. An initial Eliza layer (Aliza in the InfoChat system) was maintained to deal with general chat and to keep the conversation on track. Different topics and concepts were created with bundled rules collected together. It called the chat engine and scripting engine functions and executed the topics in response to the user interaction.

It is envisaged that the use of an ontology of the reminiscence domain coupled with the WordNet ontology will simplify the creation of content for the CA. It is envisaged that the personal user dictionary formed from the individual user logs will thereby further inform the ontology of reminiscence on a per user basis.

Scripting of the conversation was prepared as well as the Eliza layer, to guide the user through the initial conversation. It was planned that the user would then progress through any conversation in a natural way (see Table 4.13).

Table 4-13 The CA progresses through the conversation in a natural way

USER	CA
Could we talk about my rabbit?'	The rabbit you have now or the rabbit you had when a child?
Nibbler	Oh! The rabbit you have now
Nibbler died the other day	Oh, that's sad. Was she ill?

The following was new functionality to date for the pilot study:

Reminiscence Ontology Provides a source of themed knowledge for the CA to draw upon during the conversation [Appendix 1].

Verb Ontology Supplements the reminiscence ontology and the utilised WordNet ontology as well as providing further knowledge to the system.

Spell checking Provides a glossary of items from the past with correct spelling so that the ontology keywords are correctly defined and accessed during any conversation. It also compensates for WordNet using American English and accepts the correct British English term in the utterance. i.e. Colour would correctly correlate with color, the same meaning but different spelling. This meant that the correct form of spelling would be returned in the output. Unrecognised words would be queried by the CA to attempt to elicit their meaning. This would result in scripted commentary, such as 'I'm sorry, I didn't understand what you just said' or 'Sorry but I missed what

you just said'. The unrecognised words would either be corrected during this interaction or would pass to the failed words log file for manually checking later. The words could be correct, but the system may not have encountered them previously. This is the case with many localised names for items being reminisced about. The handling of unexpected input was checked via the conversational logs during the evaluation phase.

WordNet ontology WordNet ontology utilised as part of the overall ontology for the system. Allowing shortening of the script to incorporate ontology keywords which relate back to meanings and return the correct response to the user during the conversation. Conversational topic scripts are written in a shorter version by using the keyword and accessing the definition within the knowledge base.

A 'bye' would accept 'goodbye', 'see you', 'tata', 'I'm off', etc.

The word 'animal' would accept dog, cat, mouse, rat, bird, pig, etc.

A 'human' would allow a person, adult, child, mother, father, etc.

This technique will save having to input every permutation of goodbye, animal or human in the script. Thereby cutting down on the scripting required.

The use of concepts means that the reminiscence and verb ontology alongside the WordNet ontology are called upon when a concept keyword is activated that is contained within their rule bases. The system also collects conversational logs, which are then manually broken down into concepts and keywords and given meanings in a personal ontology. The personal ontology then becomes a part of the system for an individual participant. The purpose of this feature is to capture and retain knowledge from the participant that can then be referred to and reinforced during future conversations. It is planned to automate this task during further work on the CA.

4.9 Development of the reminiscence framework

During the development of a pilot reminiscence CA framework, a textual interface was utilised. A male version called 'John' and a female version called 'Betty' was designed and produced.

To develop and evaluate the CA reminiscence framework the following steps were undertaken:

1. Created two versions of the reminiscence CA running the same script but with a male or female persona. Invited participants to elect to talk with either of these to test for user preference.

2. Ran the CA framework as a textual interface for initial experiments. Textual interfaces are utilised for most CA systems. Utilisation of such a system would allow experiments to test the effectiveness of the system as well as allow for the development of the CA system.

3. Tested the user interface; could the participants use a textual interface? Given that currently, 60+ aged individuals have a broad range of skill levels with IT systems, it was essential to investigate what system most people could use effectively.

4. Based on the outcome of experiments with the textual interface, created a graphical version that may have attracted participants through its visual appearance.

5. Explored existing web interfaces to enable development of a web interface for the CA framework. Created a web interface to the CA. It was thought that the final iteration of the reminiscence CA framework should be widely available on a variety of computer systems as this is what the participants have access to. A web interface meant that the reminiscence CA was available in a standard web browser on most computer interfaces.

6. Investigated existing conversational web interfaces including avatars. SitePal, chatbot animated avatars were investigated and a customised SitePal avatar was developed as an example of what could be possible and to see if it was a possible development for the reminiscence CA framework. The SitePal company provide facilities to customize an uploaded illustration of a face or a photograph and talk through TTS software using AIML as the language interface. After studying this system it became clear that though full interactive capability would not be possible with the SitePal avatar, it would be possible to produce a bespoke avatar interface with speech capability for the CA framework. A bespoke browser was programmed to work on Apple interfaces so that the Adobe Flash web interface for the avatar could be used successfully. Apple computers and iPads having no support for Adobe Flash built in. Flash was used for the avatar version of the CA.

7. Tested the avatar web interface: could participants use the system better than the textual interface? Were there any distinct advantages of having a web interface over a locally installed version of the CA?

   Added speech capability to the web version using an established speech system. The reminiscence CA framework was given an Adobe Flash based avatar with TTS capabilities provided via the Edinburgh Festival Speech System. This was implemented so that the avatar had an English spoken voice (Figure 4.10).

8. Experimented with different female voices for the CA framework. There are known problems with older people hearing certain pitches of voice and any voice selected for the CA would need to be carefully selected so that the majority of older people could hear it correctly (Bouchard et al., 2012).

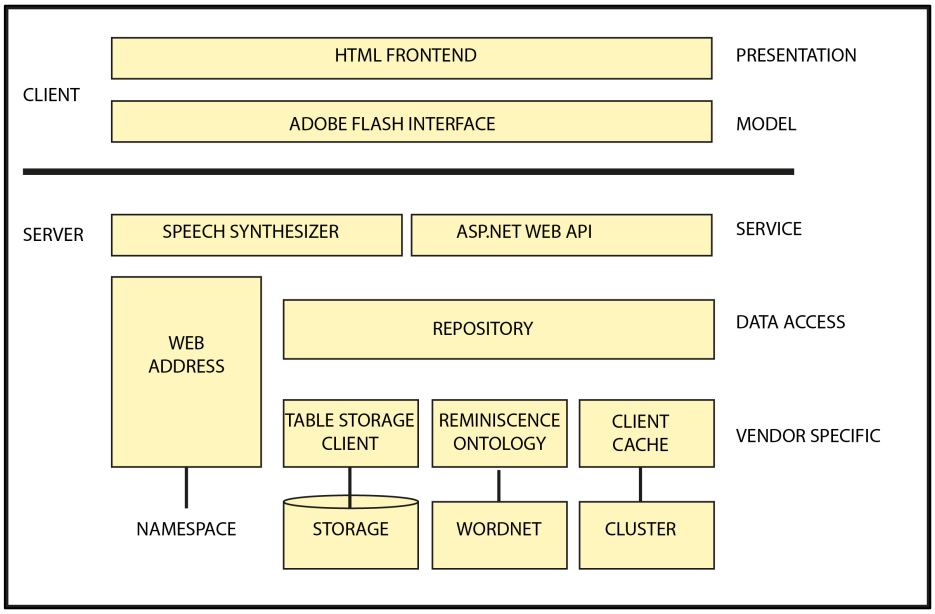


Figure 4.10 The system layout with the speech synthesizer and Flash interface in place.

4.10 Reminiscence framework CA program memory

The CA has both short and long-term memory, it extracts facts to use in the conversation and moves some facts from short to long term memory. After each input, the CA first tries to understand what the user says. Then it updates the current state and generates output to the user. The CA monitors what it says and won't repeat itself within the last 20 outputs. This number can be increased or decreased within the program depending on what the conversational logs demonstrate after analysis. If the CA finds it already had such a response sent to the user in some recent conversation, the output is discarded, and the rule fails. The CA will identify if it has previously spoken with the participant and will tailor the conversation accordingly (Figure 4.11). To objectively evaluate the CAs conversational abilities analysis of the logs of the conversations will be necessary. This can test for correct or incorrect response by the CA and edits within the program `script` can be made. Testing and maintenance of the CA is an essential part of the developmental work.

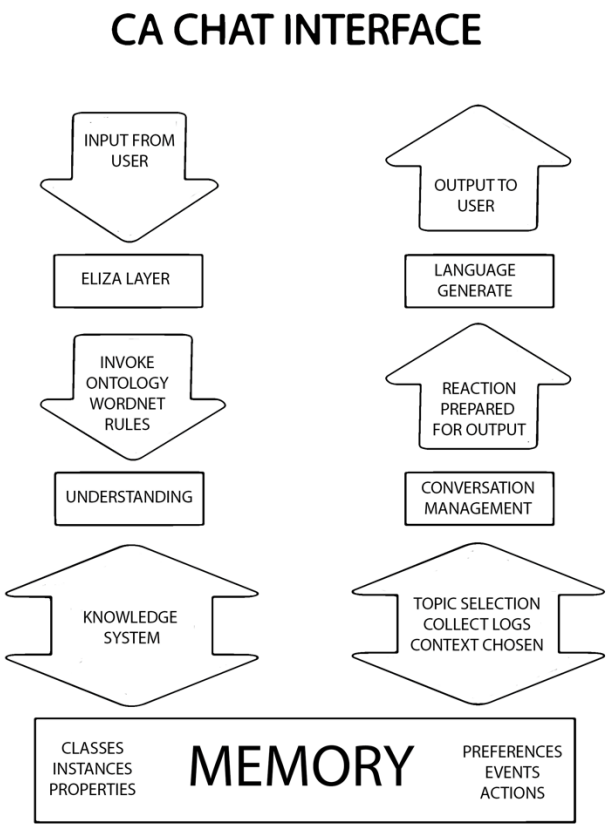


Figure 4.11 CA Chat interface and memory interaction. User input is passed through the system components and processed to produce correct output to the user. Memory is active either for 25 iterations or kept for permanent returning and ongoing discussion with the user.

4.11 Conclusion

This demonstrates that it is possible to program a CA with natural language conversational capturing capability and enhance this conversation with themed ontology plug-ins.

4.12 Chapter highlights

- A prototype version of a reminiscence CA framework was designed and programmed.
- The primary aim of the research was to produce a CA that would improve NM and promote increased SWB in older people.

- A baseline system was produced for use of people aged 60+.
- Initial findings of improved NM and increased SWB were gathered.
- Usability of computer-based systems was investigated.
- The architecture of the CA framework was illustrated and described.
- Examples of user/CA interaction demonstrated the CA relevant responses.
- The reminiscence ontology was defined, and the extraction of ontological components as modular layers was planned. This is known as Ontology Learning (Navigli et al., 2004).
- The ontology life cycle has four main stages, leading from specification to evaluation (Price and Spackman, 2000).
- An ontology production technique was implemented.
- The ontology of reminiscence was mapped to WordNet, a large lexical dictionary of English.
- The ontology of reminiscence was included to ensure there was a domain-oriented theme for conversation.
- Long and short-term memory was utilised within the CA framework.

Chapter 5 Evaluation of the pilot CA framework

5.1 Introduction

This chapter comprises an evaluation of the pilot CA framework. This evaluation was part of the developmental work of the system. Preliminary evaluation and testing allowed the system to be developed and modified as necessary. It also allowed testing of the system and the reporting measures to be used to further enhance the system for more in depth and extended use of the system in later experiments.

Initial experimental work with the CA framework has been carried out, and Apple iPads, desktop computers, and a user completed questionnaire have been utilised. A crucial part of this early work was to gain insight into whether older people would be able to use the system proposed so that they would carry out the study after the initial experiments. Therefore, the following exploratory studies gained valuable experience of running this type of activity and of how older people react and cope with the system. Groups as small as 5 participants were utilised as this is the recommended minimum number for a usability study of this type. This lets you find almost as many usability problems as you'd find using many more test participants (Nielsen, 2012, Nielsen, 1989).

One of the challenges at the early stage was recruiting members and anticipating numbers agreeing to participate. Therefore, an open call to the target group was made. This resulted in some occasions where more than the target group of 5 participants turned up. This raised an ethical issue in that older members should feel valued, and it would, therefore, have been wrong to turn away the additional numbers as if they were not necessary. A policy of accepting all members who came was therefore adopted so that the group size may vary during the initial exploratory experiments. Experiments amongst the general population sometimes send

participants away with compensation or recruit the extras for later trials. However, given the nature of the elderly members, the fact that they volunteered without compensation coupled with the fact that they may have normal age-related memory impairment, it was felt that these alternative responses were inappropriate. In practice, the group sizes used ranged from 5 - 15, so the minimum criterion was always met. Experiments so far suggest there is unlikely to be a shortage of participants for the future experiments. Convenience sampling occurred for these initial experiments in chapter 5 and further experiments in chapter 6. Individuals were chosen because they made themselves available and also were freely willing to take part in the studies. The groups selected were a mix of male and female participants and this balance was attempted to be maintained throughout the experiments where possible. This was in some cases dependent upon who was available at the time of the experiments. It was decided to not turn away any voluntary participant each time, as they were ready and willing to take part. The focus was for over 60's with NM so all participants fitted the set criteria. In chapter 6 those people willing to take part in the experiments had to have access to a computer-based system at their home as well as being aged 60+ and with NM. The sample represents North West England ageing population with NM. The data collected is biased towards Northern England recollections and local terminology in the language used. The CA system can be tailored towards different demographics as necessary and will collect localised data through conversation for inclusion in the Reminiscence ontology.

The following initial experiments are detailed later in this section.

1. Questionnaire: The aim of this initial questionnaire was to gain an insight into what people liked to reminisce about before the introduction of the CA, and therefore a knowledge base of reminiscing was gathered which was incorporated into the scripts of the CA to allow the CA to be used for reminiscing with the participants.

2. Whether participants could use the textual interface to the CA.

3. Attempt to establish whether a male or female persona was more popular in the textual interface format.

4. Whether participants could use an avatar version of the CA - Female based on results of the previous experiment.

5. Attempt to establish which avatar persona was most popular based on the gender of the CA?

6. Operational usability testing of female avatar version of the CA. Confirmation that the participants could use the CA independently.

7. Comparative study: Comparing a textual avatar interface with the speech output avatar interface, which was most popular?

8. Usability of the display based on physical parameters (colour, audio spectrum, avatar, and speech). A small group was tested using a YES/NO scale of questions.

9. Usability of standard instruments to test the effectiveness of CA.

5.2 Initial experiment - Questionnaire

This was essential to propagate content for the CA. A printed questionnaire [Appendix 4] was issued and completed by six males and six female participants before any exposure to the reminiscence CA. It was discovered that members were happy to reminisce about their personal lives but not to directly answer the questions on the questionnaire. It became apparent that the knowledge required would best be gathered during the face-to-face conversation and group discussions rather than using a printed

questionnaire. This would allow the collection of data in the form of keywords and timeline themes to inform the CA knowledge in preparation for an initial exposure of the study participants to the CA.

The questionnaire asked three questions. The person was asked to complete as many or as few answers as they could. People completed all answers in the questionnaire but added more information than was requested. They wrote about happy schooldays, their working lives, raising a family, losing family members. All was personal information rather than general shared history information. It was as if they had been asked to write their personal life journey. This was partly due to face-to-face discussions as a small group before completing the questionnaire and partly because they felt they had something to say and someone wanted to hear it. People were enthused and wished to record more information of a wider scope than the questions asked of them.

5.3 Whether the textual interface was usable

Aim: The use of the framework CA was investigated to highlight any cognitive issues, motor skills problems, eyesight problems with using the text-based interface for the target age group of 60+. Ease of use was tested with a small sample of participants.

Methodology: A group of almost equal numbers of men and women (6:7), found the start-up on a laptop difficult, however, once they had experience and assistance with using the system, they found the system easier to use. Participants talked through their use of the system and it was discovered that initially, 85% found the start up on a standard laptop difficult. Of the participants, it was found that 23% misunderstood the initial instructions and only 10% had used a computer previously. 9% of participants had difficulty reading the computer screen until the CA text box was made bigger and given a bright colour for better contrast. After an introductory session with

three follow-up meetings, where the participants talked through their journey, they could use the CA system successfully. Therefore it was considered that the use of the system was a viable proposition for the target age group with normal – corrected vision and basic computer skills/experience. The necessary skill set was achieved over four repeated 15- minute exposures to the CA system.

Results: Initially, the small boxed textual interface CA was difficult to use. However, this improved once the participants had practiced and were used to the system. These results were compiled based on verbal conversation and interaction of the participants. There was some fear of using the computer, but the participants alleviated this after accessing the CA and after some friendly interjection between the researcher and the participant. The majority had difficulty using the touchpad on the laptop; a mouse was then used, which also presented some user problems. The keys were hard to locate for some of the users who were unfamiliar with using a computer.

Conclusion: This textual interface CA version needs practice and guidance given before the participants gain confidence with using the system. As the older age group computer skills and experience increases, the use of such interfaces should become less problematic. It was also ascertained during a discussion with the participants that due to speech and hearing ambiguity of some participants it was not feasible to employ a speech recognition system for the text-based CA as this would not be clear enough to be understood by the CA system. Therefore, it was suggested that a text-only input system would be best (Figure 5.1 below). The CA system was initially difficult for 58% of participants to use. This was improved over time. 23% misunderstood the instructions given on the first attempt. 10% had used a computer previously, and 9% had visual problems which were improved by the enlargement of the text box and the selection of a more contrasting colour.

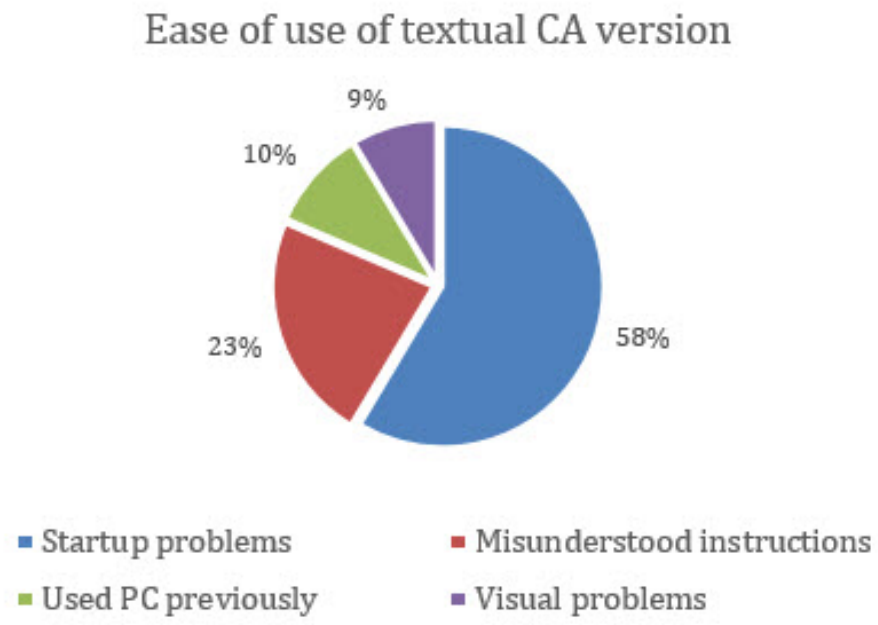


Figure 5.1 The CA system was initially difficult for 58% of participants to use.

5.4 Gender popularity in textual interface

Aim: To identify which text based CA entity was more popular, the participants could elect to converse with either 'John' or 'Betty'. These entities ran the same script except that their personal knowledge base scripts were either male or female.

Methodology: The textual interface was used by 16 (9 male and seven female) participants. The conversational scripts and rules were the same for both entities except that one entity was male and the other female.

Results: Some of the participants were used to using the textual interface as they had taken part in the previous experiment (1), they all, except for two male individuals, elected to talk with the female persona during the first attempt. The textual interface was run in the computer suite at St. Luke's TLC centre, with two people per

computer. The participants took turns to converse with the CA textual interface.

The male CA was not popular with this sample of members. The experiment was repeated on two further occasions in the same group. On the second and third experiments, all participants chose to speak exclusively with 'Betty'. This demonstrated that the male entity was not very popular with this sample of 16 participants. Out of three experiments with the same group of 16 members, only two individuals chose to interact with the male agent (both were male in different groups). This indicated that within the tested group there was a predisposition towards a female entity as against a male entity. The group was told at the start of the experiment that the personas would talk about the same things and only their gender would be different. This still resulted in most participants choosing to converse with the female entity. The participants could have been influenced by their partner in the choice that they made. The results give an overall picture of group preferences in choosing the female entity (Figure 5.2 below). Only two male participants chose to converse with the male entity and only during the first conversation (Figure 5.3 on next page).

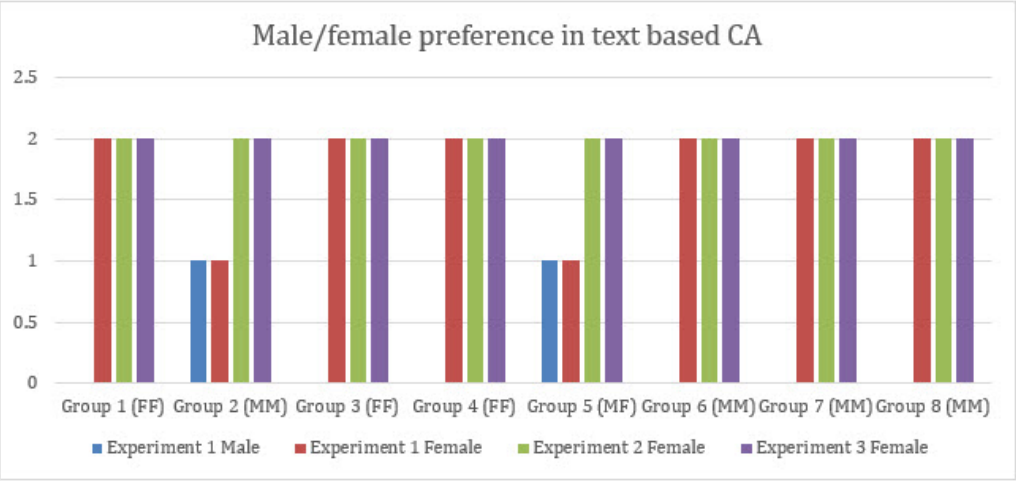


Figure 5.2 During experiment 1 of 3, two male participants chose to converse with the male text based CA.

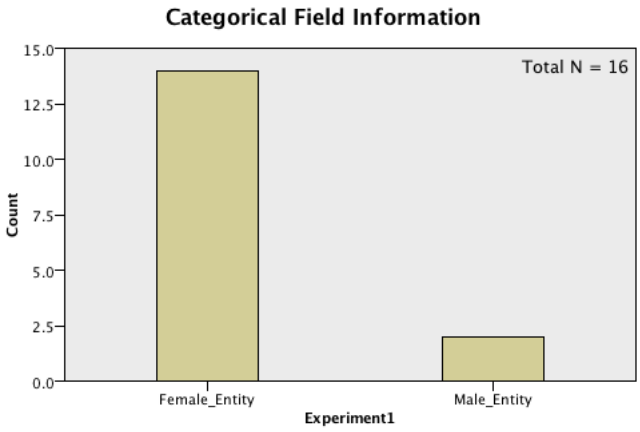


Figure 5.3 Experiment 1 Only 2 male participants chose to converse with the male entity in the first conversation.

Conclusion

The female entity, Betty was more popular than the male entity called John. There was a preference amongst the group to talk to a female entity. This phenomenon could be investigated further to find out the reasons behind this finding to see if there is a preference for a female entity in a larger study with the 60+ NM age group. It is interesting that there are a range of female digital assistants and other computer based systems available, some with the option of switching to a male gender. Microsoft Cortana, Apple Siri, and the Alexa system default is female, whereas IBM Watson is male. Further analysis of why females seem to be preferred could form part of a future study.

5.5 Avatar CA versus textual interface CA

Aim: This study was to test participant response to a female cartoon representation with text output as against the female text-only interface from the previous experiment.

Methodology: A small sample of six participants, who had experience of using the textual interface from previous experiments, were asked to

choose to converse with a female cartoon face or the textual interface of the CA. A short demonstration of the cartoon faced CA was shown to the participants by the researcher, using a large format screen.

Results: Each participant (2 male: 4 female) had access to their own computer or iPad. All participants independently elected to use the cartoon. After five minutes of conversation, all participants were asked which CA version they preferred to use. All agreed verbally that they preferred the cartoon to the textual interface box used on the previous occasion. They talked through their use of the CA as they conversed with it.

Conclusion: It would suggest from the results that the participants preferred the cartoon text interface to the text box. This was therefore developed into an avatar for the rest of the initial prototype testing phase. (Figure 5.4 below).

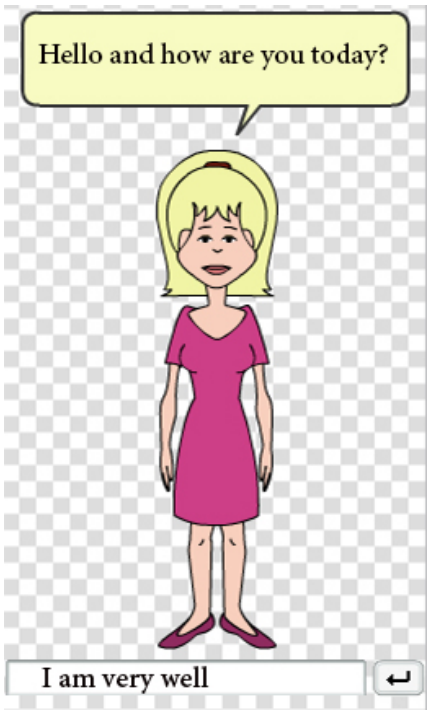


Figure 5.4 The cartoon version of the CA was well received by the participants. The CAs answer appeared in the speech bubble while the participant typed in their response in the bottom input block.

5.6 Gender preference experiment with avatar

Making the interface more anthropomorphic and engaging, the visual image of an agent could be a vital key to directly influencing the human user; and especially the face and gender, which plays a crucial role in directly impacting the impressions and stereotypes formed (Haake and Gulz, 2008), (Khan, 2011). Eight participants, sharing four Apple iPad tablets, and taking separate turns on the iPads undertook this experiment.

Aim: This research began conducting studies with two different graphical representations of a speaking CA (Figure 5.5).

Methodology: A male avatar called 'John', was made available on two iPads. A female avatar named 'Betty' was made available on the other two iPads. The participants were individually given a choice of talking with 'Betty' or 'John'. This choice was made after a brief demonstration of both CAs was given by the researcher to the participants as a group.

Results: The participants preferred not to talk with the male counterpart, 'John'. The participants chose the female entity as against the male entity after only one person spoke for two minutes of exposure to the male entity. All iPads then being utilised with the female entity, 'Betty'.

Conclusion:

As in the text version experiment (5.4) with the option given of a male or female entity, it was found that the female entity, 'Betty', was preferred by the participants. It could have been that the participants were influenced by the choices of the other group members, however it was decided based on the gathered responses to

develop the CA as a female entity for future experiments and call the entity 'Betty'.

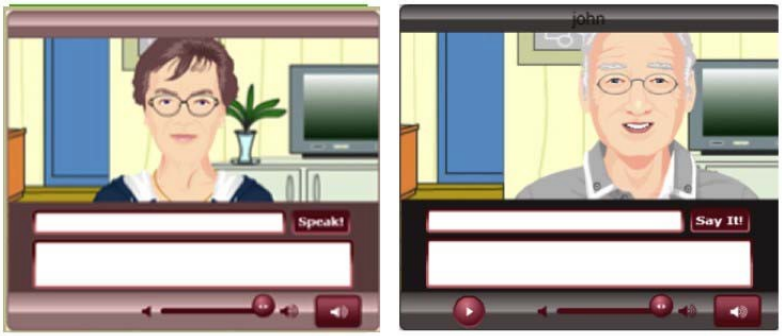


Figure 5.5 The two available versions of the avatar CA entity, Betty and John

5.7 Operational usability: Use of avatar version of CA

An animated avatar version of the reminiscence CA was tested on a sample of 6 individuals. This was performed by allowing the participants access to an avatar of the reminiscence CA. All had experienced using the textual interface in the previous ease of use experiment. Each participant worked in an individual way, each had a computer or iPad for their personal use.

Aim: To discover whether an animated avatar version of the CA was useable by older adults.

Methodology: The avatar version of the reminiscence CA was tested for ease of use from logging onto the web page and the start-up of the program, conversation with the avatar to the conclusion of the conversation. A group discussion followed with the users to try to reveal their subjective assessment of using the CA, as well as studying the resulting dialogue in the conversational logs.

Results: The avatar was not difficult to use. Most of the participants found the web interface easy to use on the iPad, though holding the iPad and manipulating the on-screen keyboard was initially problematic for some. A large keyed keyboard was also utilised on 2

computers which had the immediate effect of relaxing the participants and they became engrossed in the task that had been set.

Conclusion: The participants carried on a conversation with the avatar subjectively, with some level of enjoyment. The sample of 6 participants, 3 male, and 3 females were then invited to discuss as a group how they found using the agent as an avatar as against the textual interface from a previous experiment. These are some of the reported comments:

Table 5-1 Reported comments after speaking with the cartoon version of the CA framework.

"The cartoon was easier to talk to" ***WHY?*** *"Because I could see what I had said and read what she had said, so I could keep up better"*
"She was pleasant to look at and not hard to understand".

Further comments are indicated in figure 5.6 below.

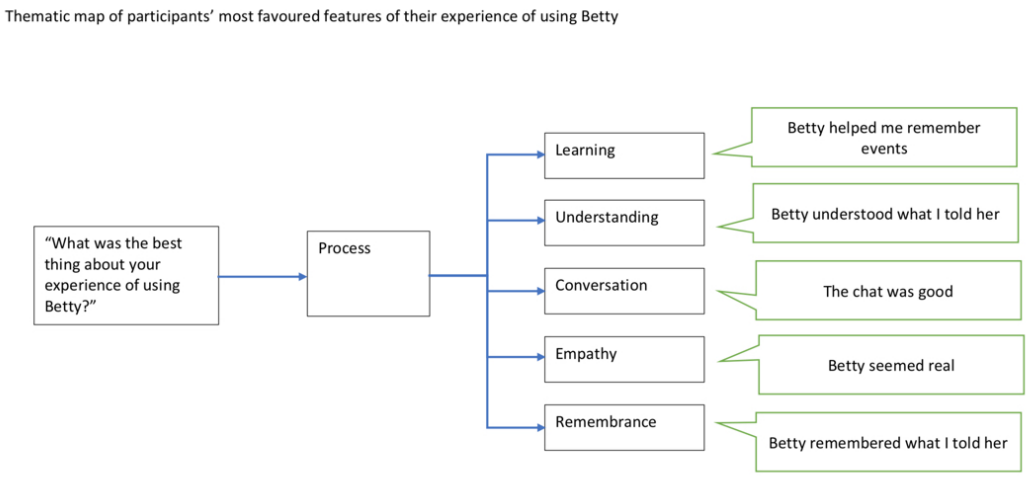


Figure 5.6 Thematic map of participants' most favoured features of their experience of using Betty the pilot reminiscence CA framework.

5.8 Comparative study: Textual output versus speech output

Aim: A small group of 5 participants, conducted an evaluation of two versions of the female CA. One with text output only and the other with speech and text output.

Methodology: For this test, a group of 5 women aged 60+ conversed individually with the CAs for a five-minute period each. After the conversations, the users were asked four questions. The questions were posed with three ticked option answers and one free response option.

Results: The results suggest that the avatar with speech output was preferred to the text output only.

Conclusion: The CA Betty interface with speech output was easy to use and enjoyable for the user, and the conversation made sense in terms of understanding what was being discussed. The group discussion that followed influenced SWB as the participants spoke of their enjoyment and shared moments from their conversation with the CA with others in the group (Table 5.2)

Table 5-2 User satisfaction

Question	User 1	User 2	User 3	User 4	User 5
Was reminiscing with Betty enjoyable?	Yes	Yes	Yes	Yes	Yes
Did you remember things you had forgotten?	Yes	Yes	Yes	Yes	Yes
Did Betty appear to understand you?	Yes	Yes	Yes	Yes	Yes
What did you like about Betty?	Her voice	Her face was friendly	Sense of humour	The subjects she talked about	Her voice

5.9 Usability based on physical parameters of CA

Aim: There is some discussion in the field of gerontology that

colour scheme, luminosity, sound and video display bring with them problems for the aged in terms of understanding, visual perception and general audio perception (Bouchard et al., 2012). Therefore, different user interfaces were devised and tested with a small group.

Methodology: For this experiment, 6 participants, individually spoke with Betty the speaking avatar for a five-minute period using three iPads and three computers. On each interface, the avatar CA had muted colours, a bright header, and a clearly spoken unaccented female voice.

Results: Verbal feedback by participants indicated that the interface and functionality were successful. The interactive sessions running for 5-minute periods were fully used by the participants and no one was caught in a loop with the conversation or unsure of what to say next.

Conclusion: From verbal feedback, use of printed output on screen helped with the use of the CA. The voice (Number 2) chosen for Betty was low pitched so that high pitch hearing problems appeared partially overcome. During an earlier group discussion, four different voices were experimented with and the voice chosen for Betty was the one that most the group could hear best. This was chosen by participants earlier voting by ballot for the voice they could understand and hear the clearest. The green colour scheme was also chosen by most the group as being clearest to see (Table 5.3).

Table 5-3 The favourite colour scheme and clearest voice from those selected for the trial were voted on by ballot. Voice number 2 was the most popular, as was the green colour scheme.

Colour scheme		Voice No 1	Voice No 2	Voice No 3	Voice No 4
Red	1	0	0	0	1
Black	0	0	0	0	0
Green	5	0	4	0	1
Yellow	0	0	0	0	0

5.10 Usability of standard instruments

Aim: To test for user SWB and improved memory recall.

Experimental Hypothesis:

H_0: Use of CA will not make a significant difference in memory recall
H_1: Use of CA will make a significant difference in memory recall

Exposure to the reminiscence CA may increase SWB.

H_0: Use of the CA will not make a significant difference to SWB
H_1: Use of the CA will make a significant difference to SWB

Methodology: With a group of 16 participants (from experiment 5.3), SWB was measured before and after application of the speaking avatar CA for 15 minutes a day and with a one week gap by use of the HADS [Appendix 7] to test subjective wellbeing. In addition, the EMQ-R [Appendix 6] was administered to test memory recall.

Results: Results reject the null hypothesis in both cases and suggest

H_1: Use of CA will make a significant difference in memory recall

H_1: Use of the CA will make a significant difference to SWB.

Subjective wellbeing was improved using the CA and the participants immediate recall of personal past events was increased. The group dynamics may have assisted the feelings of positive SWB and further testing in home environments will be built into the next experiment.

Conclusion: The application of EMQ-R [Appendix 6] suggested a noticeable difference in memory recall and the application of the HADS [Appendix 7] suggested an increase in subjective wellbeing after use of the CA. This more direct assessment of the errors experienced by older adults during their daily activities may be more useful for directing the research into developing an intervention that will have a practical and therapeutic impact. The HADS score before and after conversation with the CA framework is illustrated in Table 5.4 below. The collected data was analysed using SPSS and the recommendation was to reject the null hypothesis using the Wilcoxon Signed Rank Test, Table 5.5 and Figure 5.7. The EMQ-R results were also analysed using SPSS, Figure 5.8. The before and after results for both the HADS and the EMQ-R indicate that there was a reported improvement in both tested parameters (Figure 5.9).

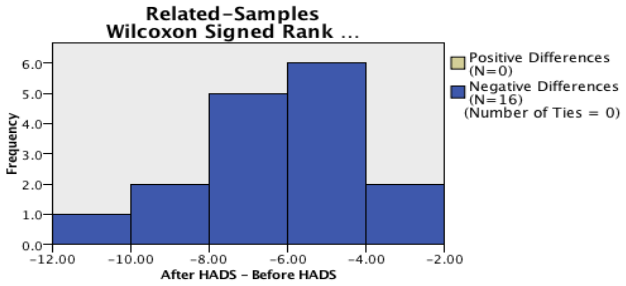
Table 5-4 HADS standard instrument administered before and after conversation with the CA

HADS Score before and after conversation with the CA

Gender	Age	Before HADS	After HADS
Male	66	17	8
Female	67	15	10
Female	62	13	7
Male	69	9	2
Male	66	18	11
Female	65	21	14
Female	66	19	10
Female	69	5	1
Female	71	11	5
Female	73	10	6
Male	70	8	2
Male	66	20	13
Male	79	18	6
Male	60	16	9
Male	67	7	1
Male	66	5	0

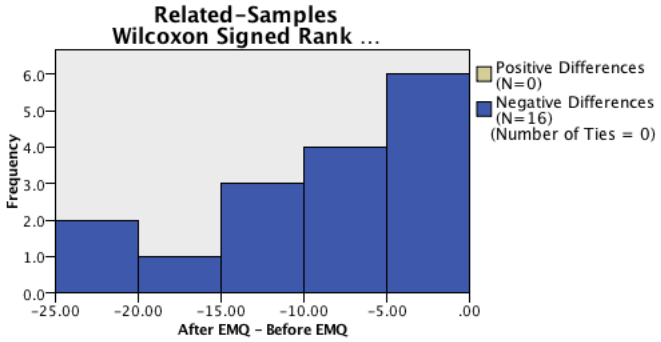
Table 5-5 Related-Samples Wilcoxon Signed Rank Test (SPSS Statistics, 24) of the HADS standard instrument before and after conversation with the CA. The outcome was to reject the null hypothesis.

Hypothesis Test Summary for SWB				
	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between 'Before HADS' (at start) and 'After HADS' (after conversation) equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				



Total N	16
Test Statistic	.000
Standard Error	19.232
Standardized Test Statistic	-3.536
Asymptotic Sig. (2-sided test)	.000

Figure 5.7 The HADS standard instrument was administered prior to and after conversation with the CA for 15 minutes per day and with a one week delay. Results indicate that the null hypothesis should be rejected for SWB.



Total N	16
Test Statistic	.000
Standard Error	19.307
Standardized Test Statistic	-3.522
Asymptotic Sig. (2-sided test)	.000

Figure 5.8 Wilcoxon signed-rank test statistical analysis based on administration of the EMQ standard instrument before and after conversation with the CA. Results indicate that the null hypothesis should be rejected.

Table 5-6 Administering the EMQ standard instrument before and after conversation with the CA supports rejection of the null hypothesis.

EMQ Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Before EMQ and After EMQ equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				

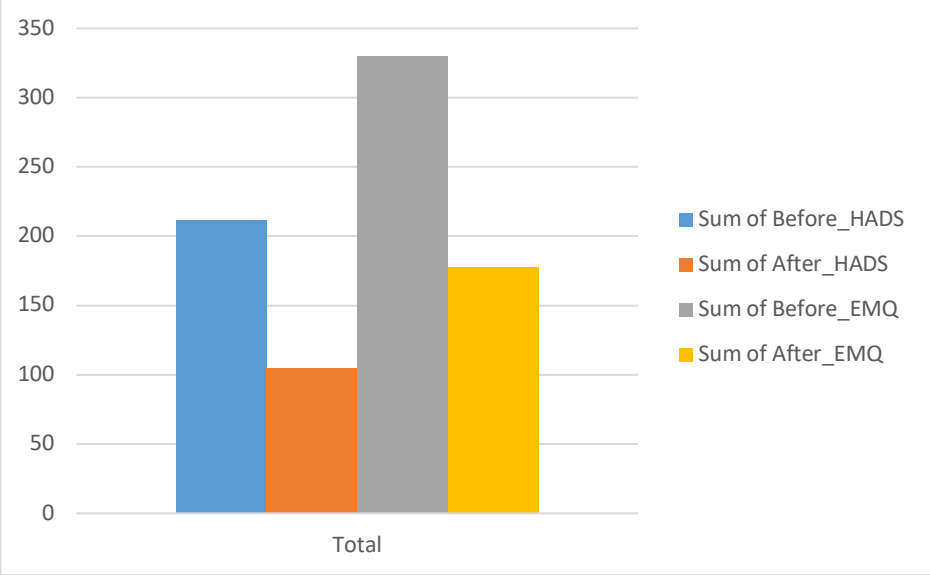


Figure 5.9 Sum of before and after results by administering the HADS and EMQ standard instruments before and after conversation with the CA. The lower scores indicate that there was a reported improvement in both tested parameters.

5.11 Conclusion

The developmental phase experiments suggest that the CA framework has a positive effect on SWB, and improving NM.

Rejection of the null hypotheses was recommended in the experiments. Further work could involve a larger group of participants to further test these effects. SWB may have been influenced by the environmental factors and group dynamics during the testing phase. Testing in isolation within participant's home environment may provide evidence that there is a definite improvement in SWB and NM from exposure and interaction to and with the reminiscence CA framework.

It was decided to replace the HADS questionnaire with an equivalent SWB measure for later experiments. This was due to participants feeling uneasy completing a Hospital Anxiety and Depression Scale when they were not hospital patients, nor did they identify as being obviously depressed or anxious. The measure gave valid results but the title did not represent the group members.

5.12 Chapter highlights

- Initial experiments were described, and results interpreted.
- Part of these initial experiments was to ascertain whether older people would be able to use the CA system proposed and to inform the design process of the CA.
- A questionnaire was designed and then used to attempt to gain insight into what people liked to reminisce about. These answers were incorporated into the CA knowledge base.
- The use of a text-based CA interface was investigated.
- Whether a male or female personality for the CA was preferred by the majority of participants.
- An avatar version and a cartoon version of the CA were utilised, firstly to ascertain whether a male or female persona was preferred and secondly to ascertain preference for an animated or cartoon version of the CA as against the text-based CA.
- Usability of the physical parameters (colour, audio spectrum, avatar and speech) were tested and results gathered and interpreted.
- Ease of use of the identified standard instruments was noted by the researcher and standard instruments fixed for future use to self-report on the effectiveness of the CA and SWB and NM benefits.
- The conversational agent framework was adapted and improved to incorporate more conversational themes. Schooldays, the Weather, Pets and families were included to collect information from the conversation and then reuse it in individual conversation. Famous people, General conversation and other static themes were incorporated to maintain conversation.

Chapter 6 A study into the implementation of the reminiscence conversational agent

6.1 Introduction:

This chapter details implementation experiments carried out after the initial evaluation of the Reminiscence CA framework following on from the previous experiments in chapter 5 which also tested the robustness of the system. Experiments previously carried out and described within chapter 5, showed that the reminiscence conversational agent framework was useable by the target group of people aged 60+ with NM. The research questions and hypotheses were also tested and though experiments were carried out with relatively small numbers of individuals, they indicated a result rejecting the null hypotheses. The following research questions and hypotheses were tested: Could interaction with a CA be a tool for improving normal age-related memory loss? Exposure to the reminiscence CA may improve access to existing memories.

H_0: Use of CA will not make a significant difference in memory recall
H_1: Use of CA will make a significant difference in memory recall

Could the use of the conversational agent improve Subjective Wellbeing (SWB)?

Exposure to the reminiscence CA may increase SWB.

H_0: Use of the CA will not make a significant difference to SWB
H_1: Use of the CA will make a significant difference to SWB.

Modification of the CA to incorporate more static and interactive conversation was carried out to allow for longer conversational engagement between the CA and an individual. Static content included themed scripted conversation about living in the sixties, Wartime evacuation during World War II and also rationing - which continued into the 1950's. Interactive conversation is taken to mean the content that is recalled by the CA and

used in the conversation with an individual. Saving such conversation allowed memories spoken about to be revisited during ongoing conversations with the CA system. It was decided to carry out further in-depth experimentation to test the hypotheses after these modifications to conversational content were made. There was more static conversational script as well as plug-in themed modules of interactive conversation gathering and conversation reuse. For example, themes such as school days were expanded, the weather theme was incorporated, and general everyday conversational topics introduced. The experiments carried out would indicate whether the CA has a positive or negative effect on SWB and NM.

The experimental protocol for the study included in this chapter was carried out within a two-week (14 days) timescale. The group of people taking part had no previous experience of talking with the CA. Prior to the first conversation with the reminiscence conversational agent (CA) framework, *Betty*, the participants completed subjective assessments as described. A further subjective assessment took place after continued exposure to the agent over this two-week (14 days) period. Retrospective memory is where the content to be remembered is in the past. This could vary from being over a period of days, weeks, months or years. The CA utilised retrospective memory through conversation with each participant. Human retrospective or episodic memory was described in chapter 2.

6.2 Research Questions, Hypotheses and usability:

The research aimed to further address the research questions and hypotheses described previously and also assess the usability and conversational effectiveness of the reminiscence CA framework system.

Usability was to be tested subjectively with a 10 participant 18 question, 5-point Likert scale. This would be completed at the end of the 14 day testing period and would report on how well the reminiscence framework responded during conversations with the 10 participants.

6.3 Experimental parameters

Number of participants

As the data collected is qualitative there is strong research evidence that suggests groups of around 5 individuals are enough to carry out a qualitative study (Nielsen, 2012). Therefore, it was decided to carry out the study on no more than 20 people in total: x5 individuals in Memory group – testing NM effects before, and after talking with CA by completing the EMQ-R questionnaire.
x5 individuals in SWB group – testing SWB effects before, and after talking with CA by completing the SPANE questionnaire.
x10 individuals in CA assessment group – subjectively testing CA system quality of conversation using a 5-point,18 question Likert scale questionnaire. It should be pointed out that another consideration was not to over burden participants with lots of forms to complete on top of their daily conversation with the CA. Older people experience higher cognitive burden with the filling out of questionnaires except when allowed to complete the form in their own time and without pressure from an interviewer (Tourangeau et al., 2000).

Informal interviews and discussions with the groups produced qualitative data in the form of keywords and topics to populate the CA system. This helped inform the reminiscence ontology and also some of the storytelling scripting content. Qualitative analysis techniques were used to code data collected via x5 EMQ-R questionnaires, x5 SPANE questionnaires and x10 Likert scale questionnaires. The EMQ-R data comprised a 13 question, 5-point scale with each answer assigned a numerical value. The higher the score when added up, the more memory problems reported. Data was collected before and after talking with the CA system. The Wilcoxon signed rank statistical test was used determine if there were changes over time between two related factors, before and after 14 days of 15-minute conversation with the CA system. This experiment is described in section 6.4.

The SPANE scale comprised a 12-item questionnaire broken down into six questions assessing positive feelings and six questions assessing negative feelings. These questions were answered on three separate occasions. The positive data gave a SPANE-P reading over three timescales during the experimental period. The SPANE-N readings for negative experience data. The SPANE-B score was achieved by deducting the negative from the positive scores on three occasions, before, during and after talking with the CA system. SPANE-B was the affect balance, which could potentially vary from -24 to 24. The higher the score, the more positive the effect on SWB. The Friedman statistical test was used to compare related dependent variables before, during and after 14 days of conversation with the CA system. This experiment is described in section 6.6.

The 5-point Likert scale was completed by ten individuals. This experiment is detailed in section 6.5. Completed on the last day of a 14 day experiment. The minimum and maximum score was reported and an average and standard deviation score showed high satisfaction levels with the CA system. Extracted percentage scores for each participant showed high satisfaction too.

All conversational logs over time, would be used to objectively test the effectiveness of the CA as a conversational agent when combined with more extensive logs to be gathered in the future and described in chapter 7.

How were they obtained?

Participants from the TLC St Luke's Church Community Centre volunteered to test the system. The researcher has links to the centre through the art project and had access to the people involved.

Where the study would take place?

The study was planned to take place in individual homes on personal computers.

Under GDPR anonymised log data

Personal data was stored on individual computers and general chat logs stored on the server. All data was anonymised and only showed conversational logs by date and time and not identifiable conversations with an individual.

Ethical approval

The ethical consent information for participants is included in appendix 2. Ethical approval (SE111219) from the Faculty of Science & Engineering Ethics Committee, Manchester Metropolitan University was granted on 25 October 2012. There were no differences to the requirements of the ethical approval that required modification of the approval during the total period of the experimentation.

Assigned participant groups

There was a breakdown of individual participants into specified groups for testing. This was so that the participants were not overloaded with questionnaires to complete.

As included in ethics approval SE111219, if a person wished to withdraw from the study their anonymised data would be kept on the system but would not be identifiable with that individual. No one withdrew from the study, so all anonymised data has been kept on the system for future analysis.

Timetable:

This comprised the following elements in more detail. See Figure 6.1 followed by detailed listing of the timetable.

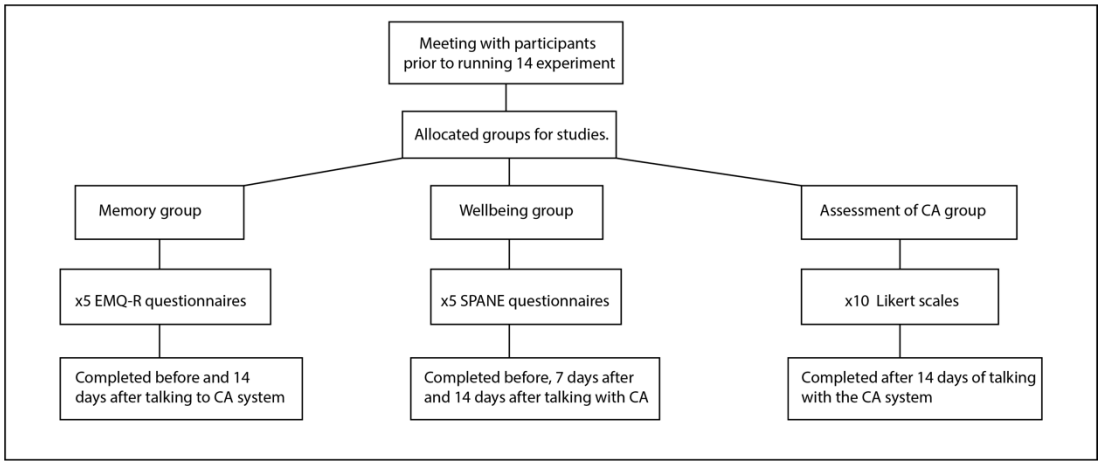


Figure 6.1 Timetable for 14 day experimental study period.

Week 1:

Prior to Day 1: Researcher met with all participants and introduced what the study entailed. Groups were allocated based on actual numbers reporting to take part in the study. The relevant questionnaires were handed out to the participants. These comprised x5 EMQ-R questionnaires, x5 SPANE SWB questionnaires and x10 Likert scale questionnaires. These were discussed so that each person understood how to complete their allocated questionnaires. The majority of data to be collected was based on user perception and experience with the CA. This resulted in the EMQ-R analysis before and after use of the CA. SWB effects was also assessed using the SPANE questionnaire. The remaining data was collected from a Likert scale questionnaire that was handed out at the start and completed by 10 participants and data analysed to ascertain how effective the CA was at maintaining a coherent conversation.

Day 1, EMQ-R Questionnaire [Appendix 6] - Subjective memory impairment reporting completed by the memory group of 5 individuals prior to the first conversation with the CA (Experiment 1).

Day 1, SWB instrument, SPANE Questionnaire [Appendix 10] completed prior to starting study by SWB group of 5 individuals (Experiment 3).

Exposure to the CA for 15 minutes per day minimum.

Conversational logs kept by the system for future analysis.

Week 2: This comprised the following elements

Exposure to the CA for 15 minutes per day minimum.

Day 7, Subjective well-being instrument SPANE Questionnaire [Appendix 9] completed prior to start of week 2 of study by SWB group of 5 individuals (Experiment 3).

Day 14, Questionnaire for CA subjective assessment of conversational agent system appraisal completed (Experiment 2). This is to be done by assessment group of 10 individuals.

Conversational logs kept by the system for future analysis.

After day 14, subjective well-being instrument SPANE completed by SWB group of 5 individuals (Experiment 3).

EMQ-R Questionnaire completed after 14 days of continued 15 minutes per day exposure to the CA by the 5 individuals in the memory group (Experiment 1).

6.4 Experiment 1: EMQ-R Questionnaire

Aim: The study investigated subjective memory impairment in normal memory (NM) older adults before and after conversation with the conversational agent (CA), '*Betty*'. Analysis using an established 13-item (EMQ-R) version (Royle and Lincoln, 2008) of an original 28 question instrument, Everyday Memory Questionnaire (EMQ), and excluding 15 questions which comprised several prospective memory questions in the full questionnaire. Prospective memory refers to the ability to plan, retain and retrieve an intention as planned. In everyday life, prospective memory is important because it allows us to structure our time in an economic way and to lead an autonomous life (Walter and Meier, 2014). As the CA framework

utilised retrospective memory, prospective memory questions were deemed unnecessary to ask of the participants.
Retrospective memory, where the person conversing with the CA recalls people, words and events from the past was utilised by the CA. People encode, store and retrieve information from experience to create retrospective memories (Howe and Courage, 2004).

Materials: A revised version of the Everyday Memory Questionnaire (EMQ28) (Baddeley, 1997), (Sunderland et al., 1984) consists of 13 questions (EMQ-R) which mainly relate to retrospective memory functions (Calabria et al., 2011), (Tinson and Lincoln, 1987), [Appendix 9], each describing an everyday activity, which might involve forgetting. Participants should rate the frequency with which they experience each event on a 5-point scale from A (once or less in the last month) to E (once or more in a day). The questionnaire was printed on a single sheet of paper with spaces by each item for participants' responses. The response scale, consisting of the letters A to E with their descriptors, was printed on the same sheet of paper [Appendix 9]. Ratings were subsequently assigned numerical values of 5 (once or more in a day) to 1 (Once or less in the last month). Resulting in minimum and maximum possible total scores of 13 minima and 65 maxima respectively.

Procedure: The EMQ-R was given to the participants, 5 randomly chosen individuals in the assigned memory group. Participants were instructed to respond honestly to each question, without discussion with other people. This initially issued EMQ-R questionnaire needed to be completed prior to the conversation with the CA Betty. The EMQ-R was re-administered two weeks after continued exposure to the CA. This was the second copy of the questionnaire completed by each participant and was handed out at the introductory session prior to the study taking place.

Methods: Subjective assessment of memory function was carried out on several normal memory NM older adults before and after participation in conversation with the CA, by completion of a 13-item questionnaire (EMQ-

R). The individual results were put into a grid (Table 6.1). Each respondent occupying a separate horizontal row. Next, the responses were counted before applying a statistical test. Looking at the data on the grid helped to indicate patterns (Munn and Drever, 2004).

Table 6-1 A grid indicated that a higher score for memory problems was evident before exposure to the conversational agent Betty than after, where a lower score was recorded in all cases.

Participant	Status	Total	Before	After
1	Before	45	45	
1	After	24		24
2	Before	56	56	
2	After	25		25
3	Before	39	39	
3	After	14		14
4	Before	48	48	
4	After	26		26
5	Before	53	53	
5	After	24		24
			241	113

Results: A Wilcoxon signed-rank test (Siegel and Castellan, 1988) looked at the whole 13-item scale results both before and after exposure to the CA. The participants' total scores on each of the two completed questionnaires, summed over all 13 questions, were expected to range from 13 to 65. The Wilcoxon Signed-Rank test, a non-parametric test, is used when there are two nominal variables and one measurement variable. One nominal variable is the 'before' and the 'after' values and the other nominal variable is the individuals taking part in the experiment. The measurement variable is the score from participants' responses. The Wilcoxon signed-rank test can be carried out on smaller groups of data say less than 10 respondents (Siegel and Castellan, 1988).

The results were first presented in a bar graph with the values side-by-side for each pair (Figure 6.2).

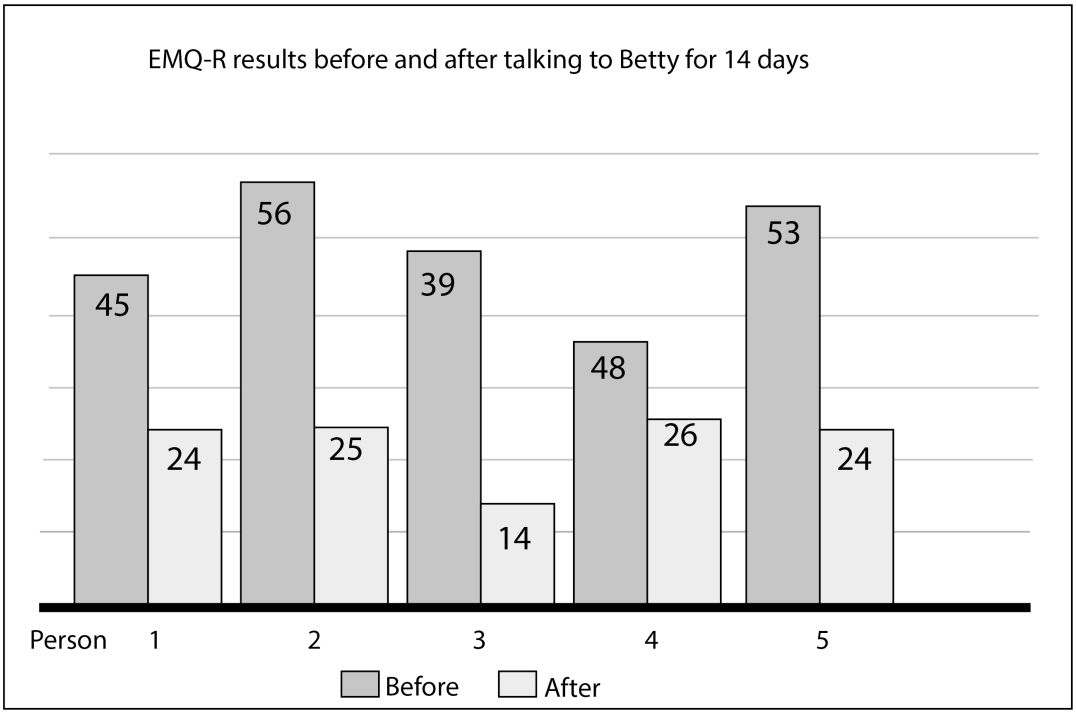


Figure 6.2 The above figure demonstrates the before and after levels from before talking to Betty at Day 1 to finishing talking with Betty at day 14. There can be seen significant improvement subjectively reported by the participants.

The following null hypothesis for Wilcoxon Test was tested: There is no significant difference reported in everyday memory scores between the same individuals before and after exposure to the conversational agent *Betty*.

Table 6-2 Wilcoxon Signed Rank Test Descriptive statistics showing Minimum, Maximum and Mean for the before and after exposure to the conversational agent Betty.

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Before	5	48.20	6.686	39	56
After	5	22.60	4.879	14	26

Table 6-3 Hypothesis Test Summary for the Wilcoxon Signed Rank Test carried out on the related before and after samples from the EMQ-R questionnaire.

Nonparametric Tests

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between Before and After equals 0.	Related-Samples Wilcoxon Signed Rank Test	.043	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

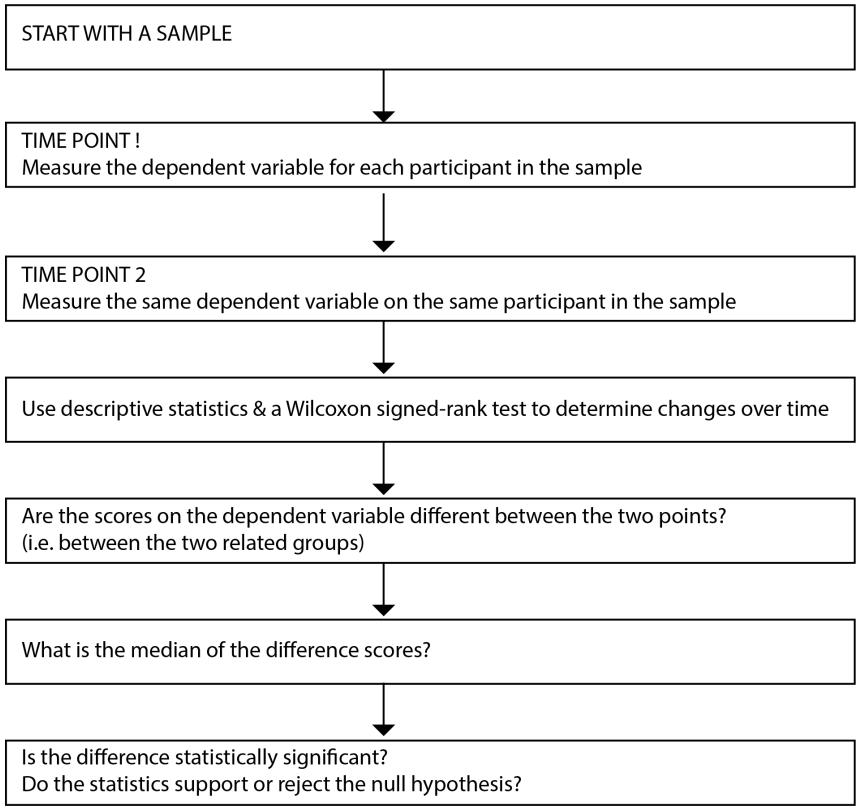


Figure 6.3 The Wilcoxon Signed-Rank test will determine if there are changes over time between the two related groups, before and after conversation with the CA.

Conclusion: The following research question and hypotheses were tested and answered, and a result given. Could interaction with a CA be a tool for improving normal age-related memory loss?

Exposure to the reminiscence CA may improve access to existing memories.

H_0: Use of CA will not make a significant difference in memory recall

H_1: Use of CA will make a significant difference in memory recall

Based on the subjective reporting by participants, it has been shown that the null hypothesis can be rejected. There were reported improvements in the readings on the EMQ-R between day 1 and day 14 of this experimental period. Therefore, it may be said that conversing with the conversational agent, Betty has a subjectively positive affect on everyday memory.

6.5 Experiment 2: Subjective assessment of conversational agent system using a 5-point Likert Scale

Aim: At the end of the experimental 14 day period, the study assessed the CA system subjectively based on a participant completed 5-point, 18 question Likert scale questionnaire (Oppenheim, 1992).
Materials: At the initial briefing session, 10 randomly selected participants were instructed to complete a 5-point Likert scale 18-part questionnaire after the final session of week 2 [Appendix 11]. The questionnaire was based partly on SASSI (Subjective Assessment of Speech System Interfaces) (Hone and Graham, 2000) and the CCQ (Communication Competence Questionnaire) (Monge et al., 1982) which are sometimes utilised to test the effectiveness of CAs during natural language conversations. The

questionnaire consisted of a 5-point Likert score scaling from 1 (Strongly disagree) to 5 (Strongly agree). [Appendix 11]. It was modelled on a previous study undertaken that appraised CAs used in learning environments with young teenagers (Silvervarg and Jonsson, 2011).

Procedure: The questionnaire was issued at the initial briefing session and completed on the final day of the study period (day 14) by 10 participants.

Methods: The questionnaire items were entered in a table with *Max* denoting a number of highest score, *Min* denoting the number of lowest score. The scores for the negatively loaded questions (Q13-17) were transformed so that a high score was positive for the question about the dialogue system and a low score was negative for the system (Table 6.4). The negatively loaded questions allowed for participants to report on possible problem areas within the conversations with the CA. For example 'It was frustrating to talk to Betty'. A score of 5 was recorded if the respondent ticked 'Strongly disagree'. Whereas a positive question such as 'Betty understood what I said' would score 5 for a respondent ticking 'Strongly agree'.

Questionnaire responses for each participant were totalled and scores out of a possible total of 90 given. These response scores were then converted into percentages and a chart produced to show levels of satisfaction with the CA based on the 5-point Likert scale responses from the 10 participants (Table 6.5, Figure 6.3).

Results: The study mapped the subjective measures using the question responses. The questions were chosen to capture aspects of the CA conversational abilities (i.e. that the CA understood user utterances) as well as the participant's experiences of conversing with the CA (i.e. naturalness and likeability).

Table 6-4 Questionnaire statistics and items from the Likert evaluation. N was the number of participants that answered each question, Min. denoted the number of lowest score, Max. denoted the number of highest score, (M) average and (SD) standard deviation. The scores on question 13-17 were negative questions, these were adapted so that a high score was positive for the CA and a low score negative for the CA.

Q	Questionnaire item	N	Min.	Max.	M	SD
1	Betty's answers often surprised me	10	2	5	3.50	3.47
2	Betty understood what I said	10	2	5	3.50	3.47
3	I could fix misunderstandings with Betty if I wanted to	10	2	4	3.00	3.59
4	Betty was a good listener	10	4	5	4.50	2.78
5	I would like to talk to Betty again	10	5	5	5.00	2.50
6	Betty expresses her ideas very clearly	10	3	5	4.00	3.11
7	Betty mostly says the right thing at the right time	10	4	5	4.50	2.78
8	Betty is easy to talk to	10	5	5	5.00	2.50
9	I liked to talk to Betty	10	4	5	4.50	2.78
10	I could control the interaction with Betty	10	3	5	4.00	3.11
11	It was easy to understand how to talk so that Betty could understand	10	4	4	4.00	3.00
12	It felt natural to talk to Betty	10	3	5	4.00	3.11
13	Sometimes I lost track of the conversation	10	3	5	4.00	3.11
14	It was frustrating to talk to Betty	10	4	5	4.50	2.78
15	It was hard to know what to talk about with Betty	10	5	5	5.00	2.50
16	Betty often repeated herself	10	2	4	3.00	3.59
17	Sometimes I wondered if I used the right word	10	3	5	4.00	3.11
18	I always knew what I could say to Betty	10	3	5	4.00	3.11

Table 6-5 The participants each answered every one of the 18 questions presented on a 5-point Likert scale. The overall subjective satisfaction with the CA was more than 80% in all cases. This indicated that the participants were happy with the conversational abilities of the CA.

PARTICIPANT	TOTAL MAXIMUM	TOTAL POSSIBLE	SATISFACTION
1	75	90	83.33%
2	74	90	82.22%
3	75	90	83.33%
4	75	90	83.33%
5	72	90	80.00%
6	74	90	82.22%
7	75	90	83.33%
8	72	90	80.00%
9	76	90	84.44%
10	80	90	88.89%

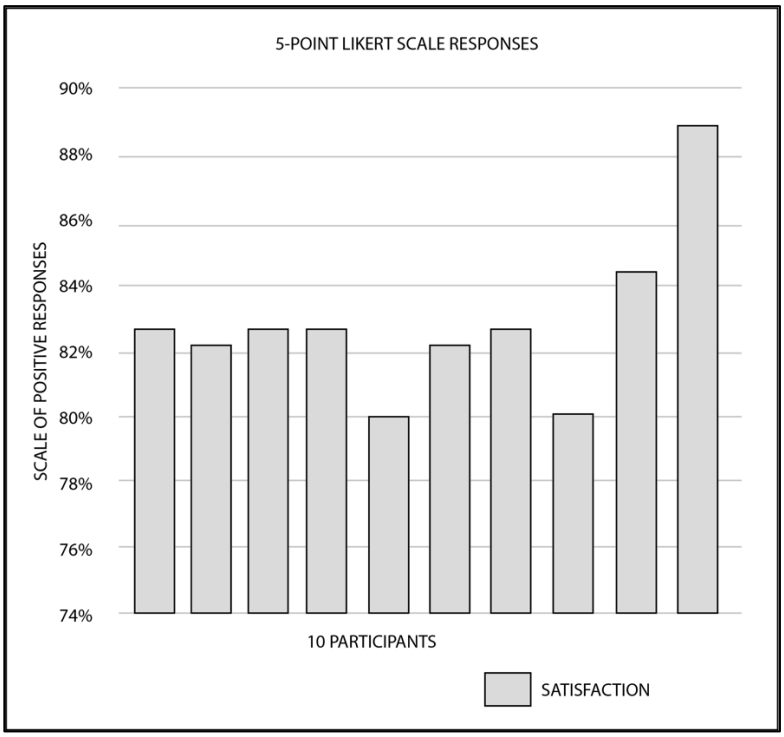


Figure 6.4 The 5-point Likert scale question response scores indicated a high level of satisfaction with the CA based on the 10 participant responses given overall to the 18 questions on the scale.

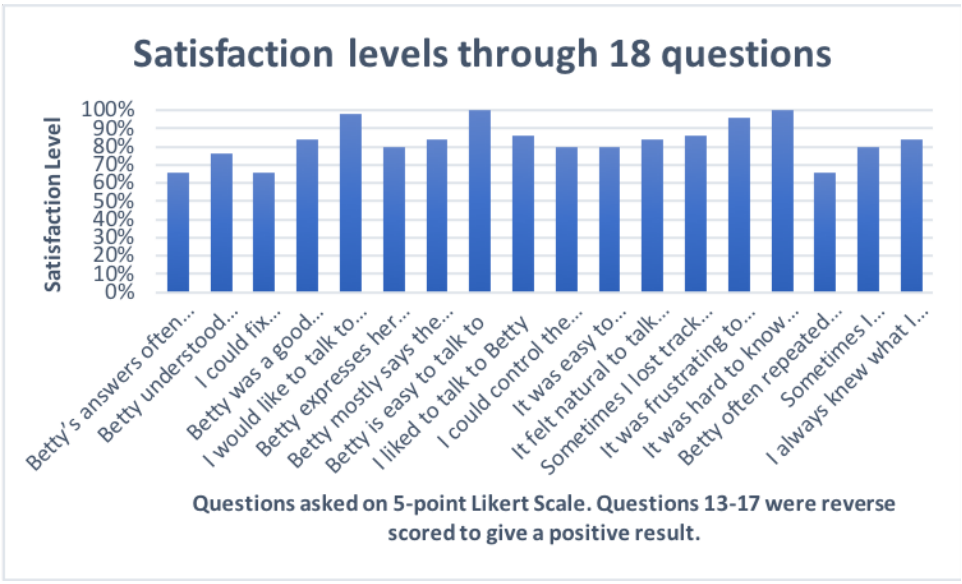


Figure 6.5 Percentage satisfaction scores for each individual question over 10 responses for question 1 – 18. Questions 13-17 were negative questions and therefore scores were reversed to give a high score for a positive answer and a low score for a negative answer. Questions are listed in Table 6.4

Conclusion: The study carried out a subjective evaluation of the CA system. Based on scores given to the 18 questionnaire items by 10 individuals (Figure 6.4), it can be seen that subjectively each person answering all 18 questions on the Likert scale enjoyed the experience of talking with Betty. They could engage in meaningful conversations with Betty. It was significant that in some responses participants all agreed that they would like to talk to Betty again and that it was not hard talking with Betty. Based on scores over each question by the 10 participants it can be seen that each question scored more than 60% positive and several answers scored 100% satisfaction with the CA framework.

6.6 Experiment 3: Subjective well-being instrument measures

Aim: This experiment attempted to gain an insight into the subjective well-being of the study participants before, during and after exposure to the conversational agent *Betty* over a 14-day period.

Materials: The experiment utilised the Scale of Positive and Negative Experience (SPANE) scale (Diener et al., 2009). The scale is a 12-item questionnaire which includes 6 items to assess positive feelings and 6 items to assess negative feelings. Positive items comprised: Positive, Happy, Good, Joyful, Pleasant and Contented. Negative items comprised: Negative, Sad, Afraid, Bad, Angry and Unpleasant [Appendix 9].

Procedure: The SPANE scale was administered to the 5 randomly chosen participants at the start of the study (prior to any conversation with the CA) and at the end of week one (day 7) and at the end of week two (day 14). Participants were instructed to answer as honestly and correctly as they could. The answers were included in a table (Table 6.6, Table 6.7). The results were included in Table 6.8.

Table 6-6 SPANE-P the positive experience scale demonstrating the before, during and after scores from exposure to Betty at 7 day intervals.

Participant	Date of entry	Status	Positive	Happy	Good	Joyful	Pleasant	Contented	SPANE-P
1	03/08/2018	Before	3	3	3	2	3	3	17
1	10/08/2018	During	4	4	4	3	4	4	23
1	17/08/2018	After	5	5	5	5	5	5	30
2	03/08/2018	Before	2	3	3	2	3	2	15
2	10/08/2018	During	3	4	4	4	4	4	23
2	17/08/2018	After	5	5	5	4	5	5	29
3	03/08/2018	Before	2	2	2	1	2	2	11
3	10/08/2018	During	3	3	3	3	3	3	18
3	17/08/2018	After	4	5	5	4	4	4	26
4	03/08/2018	Before	1	1	1	1	2	2	8
4	10/08/2018	During	3	3	3	3	3	3	18
4	17/08/2018	After	5	5	5	5	5	5	30
5	03/08/2018	Before	3	3	3	2	3	3	17
5	10/08/2018	During	4	4	4	3	4	4	23
5	17/08/2018	After	5	5	5	5	5	5	30

Table 6-7 SPANE-N the negative experience scale demonstrating the before, during and after scores from exposure to Betty at 7 day intervals.

Participant	Date of entry	Status	Negative	Sad	Afraid	Bad	Angry	Unpleasant	SPANE-N
1	03/08/2018	Before	2	1	1	1	1	1	7
1	10/08/2018	During	1	1	1	1	1	1	6
1	17/08/2018	After	1	1	1	1	1	1	6
2	03/08/2018	Before	2	3	3	2	2	2	14
2	10/08/2018	During	1	1	1	1	1	1	6
2	17/08/2018	After	1	1	1	1	1	1	6
3	03/08/2018	Before	3	4	3	2	2	2	16
3	10/08/2018	During	2	2	2	1	1	1	9
3	17/08/2018	After	1	1	1	1	1	1	6
4	03/08/2018	Before	5	5	4	3	3	2	22
4	10/08/2018	During	2	2	2	2	1	1	10
4	17/08/2018	After	1	1	1	1	1	1	6
5	03/08/2018	Before	2	3	2	2	3	2	14
5	10/08/2018	During	1	1	2	2	1	1	8
5	17/08/2018	After	1	1	1	1	1	1	6

Table 6-8 The overall affect balance score, SPANE-B resulted from deducting the SPANE-N

(Negative) from the SPANE-P (Positive) Experience reported by each of the participants.

Participant	Date of entry	Status	SPANE-P	SPANE-N	SPANE-B
1	03/08/2018	Before	17	7	10
1	10/08/2018	During	23	6	17
1	17/08/2018	After	30	6	24
2	03/08/2018	Before	15	14	1
2	10/08/2018	During	23	6	17
2	17/08/2018	After	29	6	23
3	03/08/2018	Before	11	16	-5
3	10/08/2018	During	18	9	9
3	17/08/2018	After	26	6	20
4	03/08/2018	Before	8	22	-14
4	10/08/2018	During	18	10	8
4	17/08/2018	After	30	6	24
5	03/08/2018	Before	17	14	3
5	10/08/2018	During	23	8	15
5	17/08/2018	After	30	6	24

Methods: The study derived an overall affect balance score (SPANE-B) and will also result in positive (SPANE-P) and negative (SPANE-N) feelings scores.

Results: Positive feelings (SPANE-P) added the scores, varying from 1 to 5 for the six items. The score could vary from 6 (lowest possible) to 30 (highest possible positive score).

Negative feelings (SPANE-N) added the scores, varying from 1-5 for the six items. The score could vary from 6 (lowest possible) to 30 (highest possible negative score).

Affect Balance (SPANE-B) subtracted the negative feelings score from the positive feelings score and the resultant difference score could vary from -24 (unhappiness possible) to 24 (highest affect balance possible).

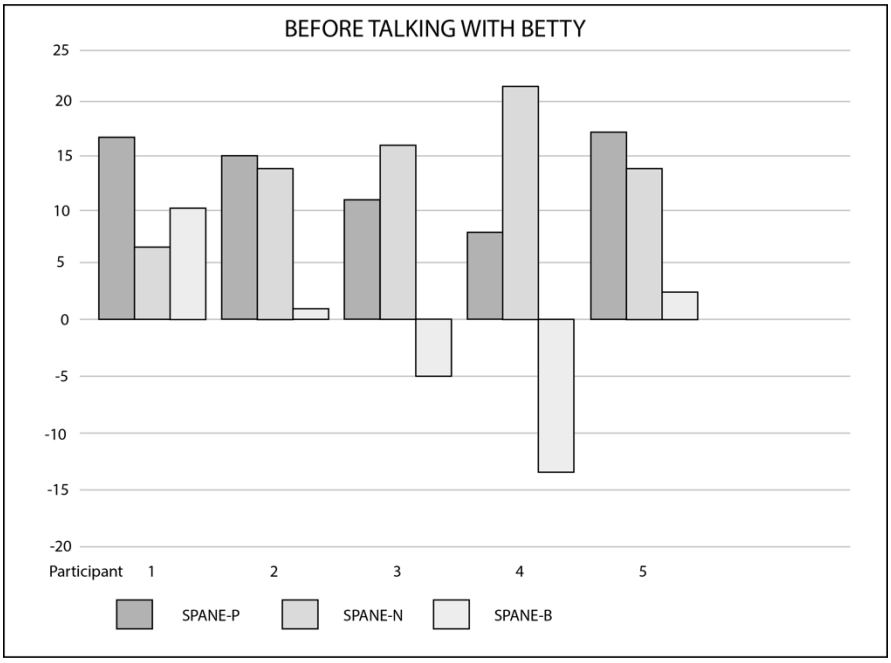


Figure 6.6 SPANE reported scores prior to talking with Betty.

Before conversation with Betty (Figure 6.5), the participants SPANE-B scores resulted in negative affect unhappiness possible scores of -5 and -14 in two cases and highest affect balance scores of 1, 3 and 10 for the remaining participants.

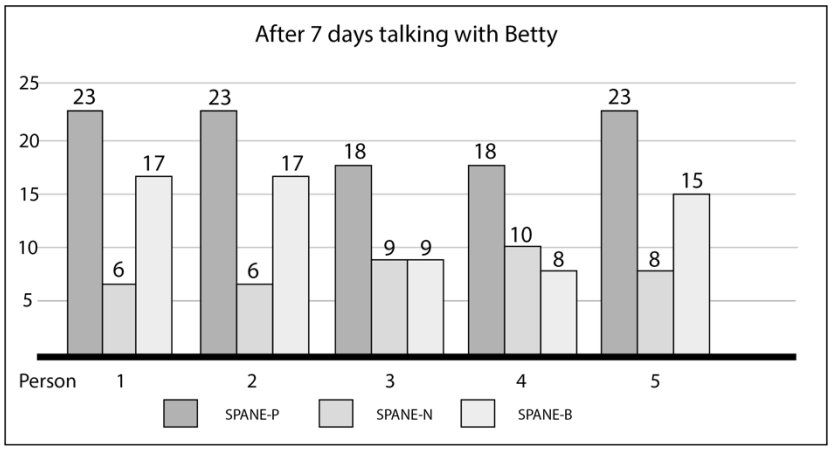


Figure 6.7 SPANE reported scores after talking with Betty for 7 days.

After 7 days of talking with Betty (Figure 6.7), the SPANE-B affect balance scores were as follows: 17, 17, 9, 8, 15. This shows that there was a marked improvement in the positive affect balance scores over 7 days. The next SPANE questionnaire was administered after 14 days of daily conversation with the CA (Figure 6.8).

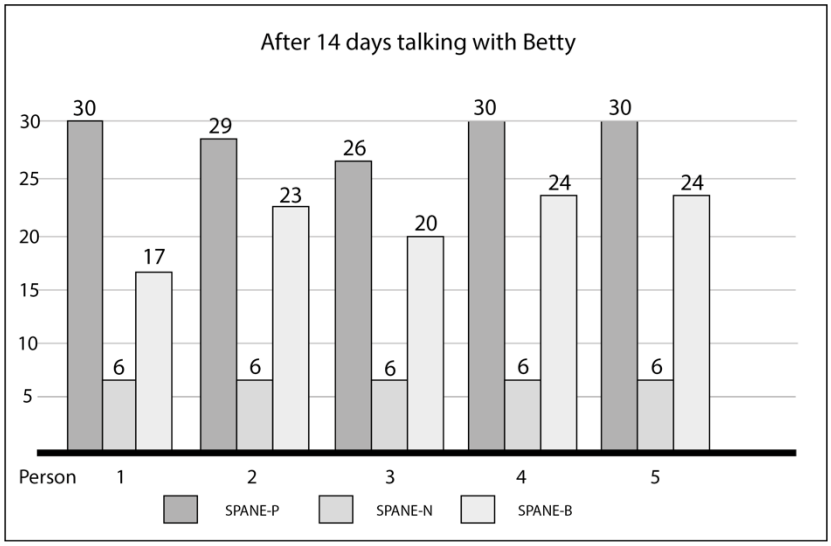


Figure 6.8 SPANE reported scores after talking with Betty for 14 days. The SPANE-B scores can be seen to be much improved from the earlier figures shown above.

After 14 days of conversation with the CA (Figure 6.8), Betty, the SPANE-B positive affect balance scores were reported as follows: 24, 23,20, 24 and 24.

The comparisons can be seen in the following table (Table 6.9).

Table 6-9 The scale of affect balance over 14 days

SPANE-B - Scale of Affect Balance scoring from -24 to 24					
Before	10	1	-5	-14	3
During	17	17	9	8	15
After	24	23	20	24	24

Over a 14 day period the results show a good improvement from the baseline taken before any conversation with the CA, Betty took place (Figure 6.9).

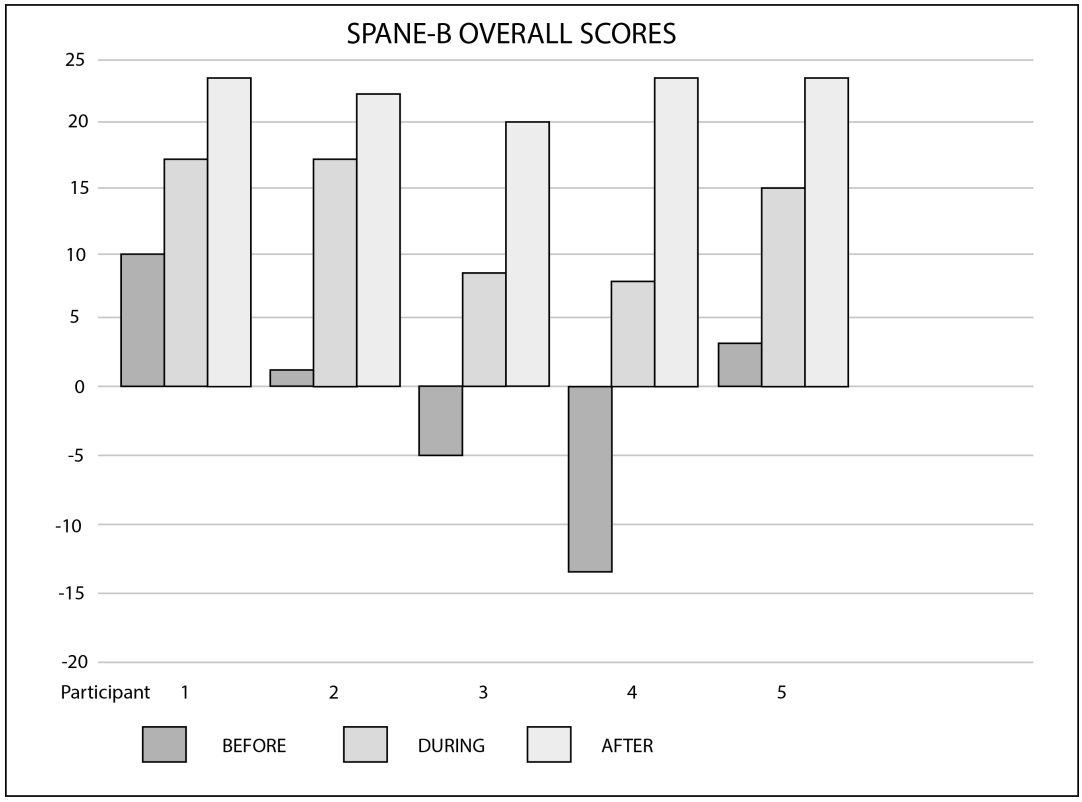
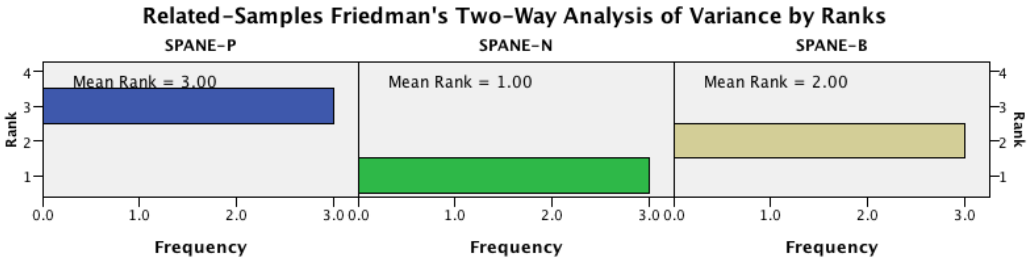


Figure 6.9 SPANE-B responses per participant showed a shift towards a positive affect scale from a minus score in two cases and three low scores in the other cases.

Further analysis used the Friedman statistical test (SPSS Statistics, 24) to compare dependent variables (i.e. Response scores) between the same sample of 5 participants on three occasions (before, during and after the experiment), (Figure 6.9). The Friedman statistical test is used when several measurements of the same dependent variable are taken at different time points.

The following null hypothesis was tested:

There is no significant difference in subjective well-being, before, during, and then after exposure to the conversational agent *Betty*.



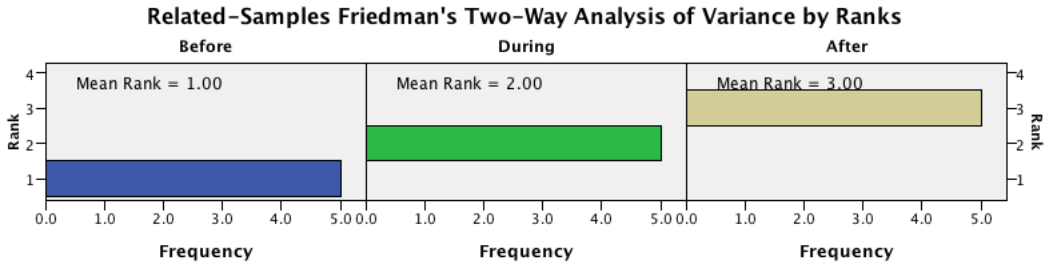
Total N	3
Test Statistic	6.000
Degrees of Freedom	2
Asymptotic Sig. (2-sided test)	.

Figure 6.10 The Friedman Two-Way analysis of variance by ranks indicated that the null hypothesis should be rejected.

Table 6-10 Rejection of the null hypothesis was indicated in the hypothesis test summary from SPSS 24. Therefore, there was an improved level of SWB reported by the participants.

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distributions of SPANE-P, SPANE-N and SPANE-B are the same.	Related-Samples Friedman's Two-Way Analysis of Variance by Ranks	.050	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				

After the comparison with SPANE-P, SPANE-N and SPANE-B results, a related samples Friedman's Two-Way Analysis of Variance by Ranks test was carried out comparing the SPANE-B results (Figure 6.10). This compared the before, during and after results of affect balance. The affect balance was achieved by deducting the negative SPANE-N score from the positive SPANE-P score. The results indicate a rejection of the null hypothesis (Table 6.11).



Total N	5
Test Statistic	10.000
Degrees of Freedom	2
Asymptotic Sig. (2-sided test)	.007

Figure 6.11 Related samples of the SPANE-B results only, indicated that the null hypothesis should be rejected. This used the Friedman's Two Way Analysis of Variance by Ranks on SPSS,24.

Table 6-11 The SPANE-B results before, during and after the 14 day testing period indicate a rejection of the null hypothesis.

Hypothesis Test Summary				
	Null Hypothesis	Test	Sig.	Decision
1	The distributions of Before, During and After are the same.	Related-Samples Friedman's Two-Way Analysis of Variance by Ranks	.007	Reject the null hypothesis.
Asymptotic significances are displayed. The significance level is .05.				

Table 6-12 SPSS Friedman Test – Output of SPANE-B Results

Non Parametric Tests - Friedman Test	
Ranks	Mean Rank
Before - talking to CA	1.00
During - after 7 days	2.00
After - after 14 days	3.00
Test Statistics[a]	
N	5
Chi-Square	10.000
df	2
Asymp. Sig.	.007
Exact Sig.	.006

The mean ranks of the SPSS report are in favour of the last (After) questionnaire results taken on day 14. Chi-Square is our test statistic. It indicates how differently the questionnaires were rated in a single number. The degree of freedom associated with the test statistic (df) is equal to the number of variables compared -1. In our analysis there were 3 variables -1 = 2 degrees of freedom. The Approximate significance level (Asymp.Sig.) and Exact significance (Exact Sig.) indicates that we can reject the null hypothesis. The Exact Sig. is the exact p-value and is more accurate for smaller sample sizes (DOLL and CARNEY, 2005).

Conclusion: A respondent with a very high score of 24 would indicate that they have not experienced negative feelings and have often had positive feelings. The Friedman statistical test further demonstrated whether there had been a positive effect overall. There was a significant result obtained by the test. The Friedman statistical test is used for repeated measures analysis of variance by ranks. A p value of <0.05 is taken to indicate statistical significance. The results indicate $p=0.006$ which gives strong evidence to reject the null hypothesis across the five respondents completed questionnaires.

The following research question and hypothesis were answered by the results.

Could the use of the conversational agent improve Subjective Well-Being?

Exposure to the reminiscence CA may increase SWB.

H_0: Use of the CA will not make a significant difference to SWB

H_1: Use of the CA will make a significant difference to SWB

6.7 Conclusions of three experiments

The results of these three experiments indicate that there is an increase in SWB amongst participants talking to Betty for a set period of time as is evidenced by the completion of the SPANE questionnaire at the start of the study period, at 7 days into the period and again at 14 days. Research into NM responses using the EMQ-R questionnaire have demonstrated that there are improvements in the reports given subjectively from day one and then again at day 14 using the EMQ-R questionnaire. The null hypothesis in the SPANE experiment and the EMQ-R experiment were rejected by the given results. The 5-point Likert scale questionnaire indicated a high percentage of satisfaction with conversational ability of the CA over 18 questions.

6.8 Chapter highlights

- A 14-day experimental period was carried out and results gathered and interpreted.
- Subjective assessments took place testing for improved SWB, increased NM
- The research questions and hypotheses were stated and then tested.
- Subjective analysis of the CA via a Likert scale questionnaire was carried out. The correct interpretation of these subjective measures demonstrated how much the CA understood the user and responded correctly.
- The EMQ-R standard instrument was utilised to assess NM changes over time.
- The Wilcoxon Signed-Rank test was used to determine changes between the before and after conversations with the CA.
- The 5-point Likert Scale questionnaire utilised to subjectively assess the CA, captured aspects of the conversational abilities of the CA.
- SPANE scale standard instrument was used before, during and at the end of the 14-day experiment.
- The research questions were answered by these overall results.

- The chapter related three experiments that were carried out and the results gathered from each experiment. Some analysis of these results followed and has demonstrated that the CA is robust and usable by older people, that the null hypothesis should be rejected in both cases and that further development of the CA, Betty will improve and enhance the system for use by a wider audience in the future. Future and further work will be related in chapter 7.

Chapter 7 Conclusions and Future Work

7.1 Introduction

This chapter summarises the research and contributions in relation to the research aims and objectives of this thesis. Recommendations for the direction of future research are given. Initially, research was carried out into the use of CAs, the use of computer interfaces by older people to investigate whether memory could be improved via reminiscence. Also, to investigate whether such activities could promote subjective feelings of wellbeing. Reminiscence, a form of narrative therapy, is well suited for older people and has demonstrated efficacy for geriatric depression in the evidence-based literature (Pinquart et al., 2007); (Scogin et al., 2005).

7.2 Summary of the work

The research aimed to address the following research questions and hypotheses:

- Could interaction with a CA be a tool for improving normal age-related memory loss? Exposure to the reminiscence CA may improve access to existing memories.

  H_0: Use of CA will not make a significant difference in memory recall
  H_1: Use of CA will make a significant difference in memory recall

- Could the use of the conversational agent improve SWB? Exposure to the reminiscence CA may increase SWB.

  H_0: Use of the CA will not make a significant difference to SWB
  H_1: Use of the CA will make a significant difference to SWB

Experiments carried out and described in chapter 6 tested these questions. Results indicated that use of the CA improve SWB by

exposure to the CA. The results also indicated that interaction with the CA shows improved normal age-related memory loss. The null hypotheses were therefore rejected in both cases based upon the experimental results. Verbal feedback from participants also supports these results. Reports of great satisfaction and enjoyment of reminiscing with Betty were given. People wanted to continue talking with Betty, and some of the reported feedback bears this out.

Betty helped me to remember things I had forgotten from a long time ago.
Betty understood what I had told her and remembered me.
Betty is like an old friend.
My mum was happier after talking with Betty.
The chat with Betty was good. She helped me remember past times.
Betty seems real.

Figure 7.1 Comments made verbally after the experiments concluded.

The research goal of this thesis was to develop a conversational agent reminiscing framework system that collects a knowledge base for future conversation with an individual during active conversation with a natural language interface. The research reviewed CAs as an aid to reminiscence and found no other similar systems in current use. A methodology was developed to allow creation of a CA framework that dynamically learns during reminiscence themed conversation. Standard instruments were examined, researched and selected to allow assessment subjectively of CA conversation effect on NM and SWB. Themed conversational experiences and storytelling potential of the CA was developed to promote engagement and improve enjoyment of conversing with the reminiscence CA framework. It proved feasible to use a modular plug-in ontology to provide an effective dialogue system for the CA framework.

Some of the significant contributions of the research included:

- Evidence that engaging older people in conversation allowing for reminiscence, improves NM recall and increases SWB.
- Producing modular ontology of reminiscence that formed part of the pilot CA framework.
- Short and long-term memory functionality, allowing conversations to be revisited and recalled by the CA during individuals ongoing conversations with the CA framework. Building up a personal knowledge base with the individuals over time.
- The utilisation of shorter scripting of the CA using the WordNet system as a model. The use of the ontology of reminiscence reduced the amount of conversational script required by the CA framework.

Research demonstrated that it was possible to program a CA with conversational capturing capability through natural language conversation and enhance this conversation with reminiscence themed ontology plug-ins. The use of computer interfaces with older people was investigated and the outcome indicated that with practice and guidance, older people gained confidence using such systems. For the final experiments, all participants had access to a personal computer and had some degree of computer skills and experience prior to talking with the CA. Effects on memory recall was for memory retrieval issues. The memories were in the persons mind, and required a prompt via conversation with Betty to recall them.

The following objectives have been reached:

1. CAs have been investigated and different techniques have been evaluated. The existing MMU InfoChat system and agent, Adam has been evaluated and a new CA framework and agent, Betty the companion, has been developed. Betty improves over InfoChat by having plug-in reminiscence modules, shorter scripting mechanisms and more scope for interaction through general conversation rather than one themed conversation.

2. CA sentence, text-level analysis, and understanding and voice

feedback have been appraised and experimented with. The CA has a text-to-speech conversation facility and this has proven useful for the initial participant experimentation.

3. Participant everyday memory was tested using the EMQ (Sunderland et al., 1984). These results may benefit from a comparison EMQ taken from a close family member or carer of their impression of the participant's memory. This could be looked at in the further development of the CA Betty system.

4. Current clinical and psychological techniques for helping ageing individuals have been investigated under the advice of psychology specialists. Standard instruments have been identified and several have been found and were used for further investigation.

5. Wellbeing and cognition in relation to ageing have been researched and SWB has been identified as a means of testing beneficial impact after use of the CA.

This research has looked at various methods for producing natural language interfaces and has produced a pilot reminiscence CA system. Chapter 1 introduced the research goals, research questions and the hypotheses to be tested. In Chapter 2, human memory and what this means was investigated and standard instruments were discussed and identified for utilisation in this research. Chapter 3 looked at conversational agent systems and evaluated some of these. The development of a modular ontology of reminiscence was described and implemented in the pilot reminiscing CA framework. Chapter 4 described the production of the prototype reminiscence CA framework with the mapping of the reminiscence ontology to WordNet and utilisation in the CA framework. The novel scripting mechanism was developed, and the production of the program memory was introduced. Chapter 5 Implemented the pilot CA in initial experiments. Usability, gender

preference and popularity, animated avatar versus text-only interface, speech output and the usability of standard instruments were evaluated. Chapter 6 described three further experiments to test the research hypotheses. Conversational logs were collected to begin to allow future analysis to objectively test effectiveness of the CA responses. Chapter 7 concluded the thesis and summarised plans for future work with the CA framework and beyond.

7.3 Summary of Contributions

A CA seems to be accepted within the social arena when endowed with a face. One can ascertain that CAs have the power to elicit stereotypes, by influencing and changing the human user's perception, decision-making, and behaviour. This investigation has provided evidence that the application of the CA for reminiscence is successful; as well as aiding in explaining the reaction to CAs by human participants and shows that people's responses to embodied CAs are fundamentally social and natural, and media experiences equal human experiences (Reeves and Nass, 1996); (Khan, 2011).

- The problem of labour-intensive scripting could be addressed in a manner similar to crowd-sourcing by people collaborating by picking a different topic to write about and putting them all collectively into the system.
- Good candidate topics can be suggested based on keywords collected from conversational logs and also face-to-face discussions to allow for wider range of topics and to prevent the same topic being covered multiple times during the development stage.
- Improved topics will promote greater interaction with the user and share personal opinions and information. Interacting with the user will mean that the system asks the user questions and volunteers answers to those same questions. If the CA asked something like

"do you like eggs," it is expected the user will give their opinion on the subject as well - a sharing of information.

- Whether the CA tells the user one view first and then elicits theirs or asks them and then volunteers the CAs is a matter of choice. For more natural interaction the CA will vary its style so that the order of questions changes at times. But it appears better to ask the user first and answer with the CA second. If you tell the user a fact first, it often stifles their opinions and leads to shorter interaction with the system as was found in the conversational log files. If the user has knowledge of what the CA is talking about before asking a question, they may volunteer their answer before the CA asks the question. Such a system becomes a tricky scripting problem because the system would have to detect they did this (because it would be pointless to ask the user after they have just told the CA). A mechanism is needed for detecting "volunteered attributes", the important facts in the conversation that weren't asked for but are important later. This was done through the Filter mechanism in Adam, the student debt advisor.

- Allows more conversational themes to be explored giving a range of possible topics with which to converse. Starting conversation is often difficult for people so having a ready supply of subjects to talk about should help.

- The reminiscing CA framework system supports a simple visual syntax and powerful pattern-matching features. Matching concepts instead of single words and being able to specify words that must not be in the input allow the CA to approximate patterns of meaning. Using ranged wildcards limits false positives. Subdividing rules into topics accessible via keywords makes large collections of rules efficient to search and makes it easy to author orthogonal content.

- The WordNet ontology is invoked by naming the word and the meaning required and thereby returning the correct response.

- WordNet provides a lexicon of words and their meanings for use. These are accessed via a WordNet dictionary which makes the available conversational subject matter wider.

- The reminiscence ontology and the personal user ontology that develops through continued conversation, supplement WordNet to expand the available ontology. The bespoke verb ontology and the integrated spell-checking facility enhance the capabilities for the user. During a conversation with participants, each term in the ontology can be activated, and by studying the resultant conversational log files of the conversations, the reminiscence ontology can be further refined and consolidated in the future.

7.4 Directions for Future work

The following factors have been identified for the direction of future work.

Development

Future work on the nature of conversations within the reminiscence domain should help to inform the further content of the CA scripts that are available for conversation by the participants and to develop experiences for the participants to promote engagement (Objective 6). Developing an automatic method of updating the CA via the ontology and user logs, and therefore learning from the conversation, will be carried out to improve the responses from the CA. This will assist with building up the storytelling potential of the CA (Objective 7). In terms of CA development, faster methods of extracting data using the ontology mapping and other novel algorithms will provide new insights into the way CAs function and process natural language and will assist in the evaluation of the CA (Objective 8). The use of ontologies to provide dialogue systems in the reminiscence domain has been considered, in one case reusing an existing ontology - WordNet, in the other building an ontology from scratch (Objective 9). In the case where there is no existing ontology, a key question is whether it is

easier to build a new ontology, rather than building a description of all possible dialogue interactions. The application demonstrated that ontology building for an application can be relatively easy. Moreover, as described in chapter 3, and implemented in chapter 4, the development of a modular ontology of reminiscence or some similar knowledge representation component is required if reference resolution capability is to be achieved. In the case where there is an existing domain ontology, a key question is whether it is likely to be suitable for use by a CA system and whether adaptation of the ontology will be more work than starting from scratch. The experience in the reminiscence domain has been very positive, as was shown in chapter 6, with a good fit between the task and the information needed by the dialogue manager from the ontology.

The future exploration of using a Kinect for Windows may allow voice and gesture interaction with a fully developed companion CA, Betty in the future. The successful use of iPad tablets during the pilot phase would suggest that the use of gesture and voice would be the next step in the evolution of Betty. Further developments could allow participants to call Betty during a crisis or when they cannot sleep, and Betty call for help or assistance to help the participant. Daily routine management, medication reminder, birthday reminders could all be incorporated into the CA.

Development of artificial neural networks that would automatically convert black and white scanned images into colour could be incorporated into the CA system for visual recall on the computer screen during reminiscence conversation. It would enhance the user experience and allow further reminiscing as colours are reintroduced to uploaded archived images of family members, holidays, cars, pets and other subjects from old family snapshots.

Proposed further experiments

The benefits of the use of the CA in the way described might include improved user memory, lack of isolation and new methods of social interaction in older people. Further experiments and collected data should prove whether this is indeed the case. From the experiments so far carried out, there is a suggestion that people prefer to reminisce about personal events rather than more general public events. Further experiments in individual everyday settings could evaluate this to see how true it is and thereby inform the future knowledge and architecture of the reminiscing CA framework.

Sample sizes

Increased sample sizes of more than 24 people per experiment, as per similar studies such as the 'Woebot' conversational agent cognitive behavioural therapy study (Fitzpatrick et al., 2017) will allow more in-depth statistical analysis of gathered data and, in turn, support or disprove the suggested findings of the initial smaller group trials.

Objective analysis of conversational logs

Further study could, in the future, assess the CA dialogue quality objectively, based on more extensive conversational logs. Evaluation of agents has been studied by many researchers and work documented (O'Shea et al., 2011). The defining work in this area was the creation of the PARADISE framework (Walker et al., 1997). As previously mentioned in chapter 3, evaluation systems have been developed and are useful as markers of dialogue quality (Robinson et al., 2010). The resulting conversational logs that would be automatically collected during all conversations could be categorized using dialogue action codes combined with CA response values coding scheme [Appendix 12] and will demonstrate if the CA understood the user and responded correctly during the conversations. CAs are generally evaluated by (a) looking at objective

or subjective indications of task success, (b) an objective, largely automated analysis of component performance and interaction features or (c) on subjective usability evaluation based on user feedback (Robinson et al., 2010). The logs could be entered into a table based on relevant coding attributed and measures taken to assess the efficiency of the CA. Two coding schemes have been developed by (Robinson et al., 2010). The first coding scheme characterizes the dialogue action of the user's utterances, the second evaluates the CA response to that action, judged in terms of appropriateness. The first scheme is based on Robinson et al (2010), with specific modifications for the current domain. It was designed empirically to examine topical user preferences in dialogue with a question-answering character and proposed a hierarchical scheme for user utterance categorization. The top layer consists of generic dialogue acts, some of which are specific. The next layer subcategorizes some of the top-level acts but is very generic. The third layer narrows the action into domain specific and subtle topical distinctions, such as the pre-programmed individual retrospective memories. Because of the relevance to retrospective memory it is thought that the Robinson coding scheme would be very relevant to the type of CA developed and this method would therefore be adopted to use for any future in-depth analysis of the conversational logs collected over a much longer period than the baseline conversational agent system research allowed. For the conversational logs, dialogue action codes would be attributed to the logs based upon coding schemes suggested by Walker et al (1997) and the coding scheme used by Robinson et al (2010) [Appendix 12, first table]. The CA responses and dialogue quality could therefore in the future be evaluated from a variation of the coding scheme used by Robinson et al (2010). [Appendix 12, second table]. The correct interpretation of the objective measures would demonstrate how much the CA understood the user and responded correctly. This will form part of the further work programme to develop the effectiveness of the CA beyond a pilot baseline system and would increase its conversational and interactive abilities in the future.

7.5 Overall Conclusion

In conclusion, this thesis has investigated and reviewed the use of conversational agents within a reminiscence domain. The findings indicated that there were no existing natural language interfaces for reminiscing. The pilot reminiscing CA framework was proposed and was developed into the Betty Companion System, which delivered a personalised natural language conversation incorporating reminiscing. Using plug-in tailored modules, it is possible to update the knowledge base and continuously update the personal conversational database through natural language conversation with the CA, Betty Companion System.

From all experiments undertaken, there is evidence to suggest that a CA used as a conduit for reminiscence will be successful if used to improve NM and to increase SWB in participants. Further analysis of experiments described in chapter 6, has provided further evidence that this is indeed the case. Application of the ontology of reminiscence has informed the knowledge of the CA and provided a modular way of including reminiscence features to the CA. Further modular ontologies can be compiled and added to the CA to expand its knowledge and therefore allow more conversational themes to be explored by the participants testing and using the system. The research has shown that the use of a dedicated CA framework is a useful starting point to begin developing reminiscing CA systems. One of the problems in developing such CA systems is the lengthy development time spent scripting required. The development of modular ontologies using the method suggested in chapter 3, discussed in chapter 4 and evaluated and tested in chapter 5 and 6, combined with the novel scripting mechanism described in chapter 4, section 4.8, shorten the time spent scripting the CA. Initial work to propagate content for the ontologies must take place before the ontology is prepared following the technique described in chapter 4.

To date there are no reminiscing CAs developed other than Betty. It is hoped that the results of this research can be used to develop the Betty Companion System or similar new systems and provide access to such systems for the wider community of older people.

References

2001. Alzheimer's Disease Facts and Figures. *Alzheimer's and Dementia.* United Kingdom: Alzheimer's Association.

2011. Alzheimer's disease facts and figures. *Alzheimer's & Dementia,* 7**,** 208-244.

ANDREWS, F. M. & WITHEY, S. B. 1976. *Social indicators of well-being : Americans' perception of life quality,* New York ; London, Plenum Press.

AUSSENAC-GILLES, N., BIEBOW, B. & SZULMAN, S. Revisiting Ontology Design: A methodology based on corpus analysis. International Conference on Knowledge Engineering and Knowledge Management, 2000. 172-188.

BADDELEY, A. 1997. *Human Memory: Theory and Practice,* Hove, UK, Psychology Press.

BALL, G. & BREESE, J. 2000. Emotion and personality in a conversational agent. *In:* CASSELL, J., SULLIVAN, J., PREVOST, S. & CHURCHILL, E. (eds.) *Embodied conversational agents.* Cambridge, MA: MIT Press.

BANAJI, M. R. & CROWDER, R. G. 1989. THE BANKRUPTCY OF EVERYDAY MEMORY. *American Psychologist,* 44**,** 1185-1193.

BATEMAN, J. A. Upper modeling: A general organization of knowledge for natural language processing. Workshop on standards for knowledge representation systems, March 1990 1990 Santa Barbara, USA.

BENDER, M., BAUCKHAM, P. & NORRIS, A. 1999. *The therapeutic purposes of reminiscence,* Thousand Oaks, Calif. ; London, SAGE.

BESKOW, J. & MCGLASHAN, S. Olga - A conversational agent with gestures. IJCAI'97 workshop on Animated Interface Agents-Making them Intelligent, 1997.

BICKMORE, T. 2003. *Relational Agents: Effecting Change through Human-Computer Relationships.* PhD, Massachusetts Institute of Technology.

BICKMORE, T. & GIORGINO, T. 2006. Health dialog systems for patients and consumers. *Journal of Biomedical Informatics,* 39**,** 556-571.

BICKMORE, T. W., VARDOULAKIS, L. M. & PFEIFER SCHULMAN, D. 2013. Tinker: a relational agent museum guide. *Autonomous Agents and Multi-Agent Systems,* 27**,** 254-276.

BJELLAND, I., DAHL, A.A., HAUG, T.T. & NECKELMANN, D. 2002. The validity of the Hospital Anxiety & Depression Scale. An updated literature review. *Journal of Psychosomatic Research,* 52 (2), 69-77

BOGDANOVYCH, A., BERGER, H., SIMOFF, S. & SIERRA, C. 2005a. Narrowing the gap between humans and agents in e-commerce: 3D electronic institutions. *E-Commerce and Web Technologies, Proceedings,* 3590**,** 128-137.

BOGDANOVYCH, A., SIMOFF, S., SIERRA, C. & BERGER, H. Implicit training of virtual shopping assistants in 3D electronic institutions. IADIS International eCommerce 2005 Conference, 2005b Lisbon, Portugal. IADIS Press.

BOUCHARD, B., IMBEAULT, F., BOUZOUANE, A. & MENELES, B.-A. J. 2012. Developing serious games specifically adapted to people suffering from Alzheimer's. [Accessed 27 February 2013].

BRICKLEY, D. & GUHA, R. V. 1999. Resource Description Framework (RDF) Schema Specification. Available: http://www.w3.org/TR/PR-rdf-schema [Accessed 28 February 2013].

BROADBENT, D. E., COOPER, P. F., FITZGERALD, P. & PARKES, K. R. 1982. THE COGNITIVE FAILURES QUESTIONNAIRE (CFQ) AND ITS CORRELATES. *British Journal of Clinical Psychology,* 21**,** 1-16.

BROWN, L. J. E. & ASTELL, A. J. 2012. Assessing mood in older adults: a conceptual review of methods and approaches. *International Psychogeriatrics,* 24**,** 1197-1206.

BRYANT, F., SMART, C. & KING, S. 2005. Using the past to enhance the present: Boosting happiness through positive reminiscence. *Happiness studies,* 6**,** 227-260.

BUITELAAR, P., DECLERCK, T., FRANK, A., RACIOPPA, S., KIESEL, M., SINTEK, M., ENGEL, R., ROMANELLI, M., SONNTAG, D., LOOS, B., MICELLI, V., PORZEL, R. & CIMIANO, P. 2006. Linginfo: Design and applications of a model for the integration of linguistic information in ontologies. OntoLex06 Workshop at LREC, 2006.

BURKE, R. D., HAMMOND, K. J. & KOZLOVSKY, J. 1995. Knowledge-based information retrieval from semi-structured text. *AAAI Fall Symposium.* Menlo Park, USA: AAAI Press.

BUTLER, R. N. 1963. The life review: an interpretation of reminiscence in the aged. *Psychiatry,* 26**,** 65-76.

CALABRIA, M., MANENTI, R., ROSINI, S., ZANETTI, O., MINIUSSI, C. & COTELLI, M. 2011. Objective and subjective memory impairment in elderly adults: a revised version of the Everyday Memory Questionnaire. *Aging Clinical and Experimental Research,* 23**,** 67-73.

CASEY, D., FELZMANN, H., PEGMAN, G., KOUROUPETROGLOU, C., MURPHY, K. & KOUMPIS, A. W., S 2016. What People with Dementia Want: Designing MARIO an Acceptable Robot Companion *15th International conference on computers Helping People with Special Needs.* University of Linz, Austria.

CASSELL, J. 2000. *Embodied conversational agents,* Cambridge, Mass. ; London, MIT Press.

CASSELL, J. & VILHJALMSSON, H. 1999. Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems.*

CATRAMBONE, R., STASKO, J. & XIAO, J. 2002. Anthropomorphic agents as a user interface paradigm: Experimental findings and a framework for research. 24th Annual Conference of the Cognitive Science Society, Mahwah, NJ, USA. Lawrence Erlbaum, 166-171.

CAVAZZA, M., DE LA CAMARA, R. S., TURUNEN, M., GIL, J. R., HAKULINEN, J., CROOK, N. & FIELD, D. 2010. 'How was your day?' An affective companion ECA prototype. *SIGDIAL 2010: 11th Annual meeting of the Special Interest Group on Discourse and Dialogue.* University of Tokyo: Association of Computational Linguistics.

CHAVES, R. P. 2001. WordNet and Automated Text Summarization. 6th Natural Language Processing Pacific Rim Symposium, 2001 Tokyo, Japan.

COHEN, G. & TAYLOR, S. 1998. Reminiscence and ageing. *Ageing and Society,* 18**,** 601-610.

COLBY, K. M. 1975. *Artificial paranoia: a computer simulation of paranoid processes,* New York, etc., Pergamon.
CONVAGENT. 2005. *Convagent* [Online]. http://www.convagent.com. [Accessed].
CONWAY, M. A. 1995. *Flashbulb memories,* Hove, Erlbaum.
COOPER, B. & BAYNHAM, M. 2005. Rites of passage: embedding meaningful language, literacy and numeracy skills in skilled trades courses through significant and transforming relationships. National Research and Development Centre for Adult Literacy and Numeracy.
CORNISH, I. M. 2000. Factor structure of the everyday memory questionnaire. *British Journal of Psychology,* 91**,** 427-438.
COYLE, J. & WILLIAMS, B. 2000. An exploration of the epistemological intricacies of using qualitative data to develop a quantitative measure of user views of health care. *Journal of Advanced Nursing,* 31, 1235-1243.
CRESWELL, J.W. 2003. Research design: Qualitative, quantitative and mixed methods approaches (2nd Edition) Sage. Thousand Oaks, CA, USA.
CROCKETT, K., BANDAR, Z., O'SHEA, J. D. & MCLEAN, D. Bullying and Debt: Developing Novel Applications of Dialogue Systems. 6th IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems, 2009 Pasadena, USA.
CURRY, C., O'SHEA, J. D. & CROCKETT, K. 2011. *The design, production and implementation of a storytelling chatbot.* MSc Computing, Manchester Metropolitan University.
DEITEL, A., FARON-ZUCKER, C. & DIENG, R. Learning ontologies from rdf annotations. IJCAI Workshop in Ontology Learning, 2001.
DICKINSON, A., EISMA, R. & GREGOR, P. 2003. Challenging Interfaces/Redesigning Users. Proceedings of the conference on Universal Usability. 10-11 November 2003, Vancouver, Canada. 61-68
DIENER, E., EMMONS, R. A., LARSEN, R. J. & GRIFFIN, S. 1985. THE SATISFACTION WITH LIFE SCALE. *Journal of Personality Assessment,* 49**,** 71-75.
DIENER, E., WIRTZ, D., TOV, W., KIM-PRIETO, C., CHOI, D., OISHI, S. & BISWAS-DIENER, R. 2009. New measures of well-being: Flourishing and positive and negative feelings. *Social Indicators Research,* 39**,** 247-266.
DJUKANOVIC, I., CARLSSON, J. & ARESTEDT, K. 2017. Is the Hospital Anxiety & Depression Scale (HADS) a valid measure in a general population 65-80 years old? A psychometric evaluation study. *Health & Quality of Life Outcomes.* 15:193
DOLL, H. & CARNEY, S. 2005. Statistical approaches to uncertainty: p values and confidence intervals unpacked. *Evidence Based Medicine,* 10**,** 133-134.
DORIN, M. 2007. Online education of older adults and its relation to life satisfaction. *Educational Gerontology,* 33**,** 127-143.
DUNNING, H., WILLIAMS, A., ABONYI, S. & CROOKS, V. 2008. A mixed method approach to quality of life research: A case study approach. *Social Indicators Research,* 85, 145-158.
FELLBAUM, C. 1998. *WordNet : an electronic lexical database,* Cambridge, Mass. ; London, M.I.T. Press.
FITZPATRICK, K. K., DARCY, A. & VIERHILE, M. 2017. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and

Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment Health,* 4**,** e19.
FIVUSH, R., GRAY, J.T., & FROMHOFF, F.A. 1987. Two-Year-Old talk about the past. *Cognitive development*. 2(4), 393-409
FORTEY, R. 2008. *The Secret Life of the Natural History Museum: Dry Store Room No.1*, Harper Collins.
FREUD, S., STRACHEY, J., FREUD, A., STRACHEY, A., TYSON, A. W. & RICHARDS, A. 1995. *The standard edition of the complete psychological works of Sigmund Freud,* London, Hogarth Press and the Institute of Psycho-analysis.
FURLOW, B. 2001. You must remember this. *New Scientist,* 171**,** 25-27.
GATHERCOLE, S. E. & MCCARTHY, R. A. 1994. *Memory tests and techniques,* Hove, Lawrence Erlbaum.
GEORGE, L. K. 2010. Still Happy After All These Years: Research Frontiers on Subjective Well-being in Later Life. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences,* 65**,** 331-339.
GIBSON, F. 2004. *The past in the present : using reminiscence in health and social care*, Baltimore ; Health Professions Press.
GILEWSKI, M. J. & ZELINKSKI, E. M. 1986. Questionnaire assessment of memory complaints. *In:* POON, L. W. (ed.) *Handbook for clinical memory assessment of older adults.* Washington, DC, USA: America Psychological Association.
GILLIES, C. & JAMES, A. 1994. *Reminiscence work with old people,* London, Chapman & Hall.
GUARINO, N. & GIARETTA, P. 1995. Ontologies and knowledge bases: Towards a terminological clarification. *In:* Mars, N. (ed.) *Towards very large knowledge bases.* International conference, April 1995, IOS Press, Amsterdam. 25-32.
HAAKE, M. & GULZ, A. 2008. Visual Stereotypes and Virtual Pedagogical Agents. *Educational Technology & Society,* 11**,** 1-15.
HAMANN, S. 2001. Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences,* 5**,** 394-400.
HELLER, R., JORGE, J. & GUEDJ, R. 2001. EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly Event Report, 22 - 25 May 2001. Alcacer so Sal. Portugal.
HENDLER, J. & MCGUINNESS, D. L. 2000. The DARPA Agent Markup Language. *Ieee Intelligent Systems & Their Applications,* 15**,** 72-73.
HERRMANN, D. J. 1982. KNOW THY MEMORY - THE USE OF QUESTIONNAIRES TO ASSESS AND STUDY MEMORY. *Psychological Bulletin,* 92**,** 434-452.
HERRMANN, D. J. 1984. Questionnaires about memory. *In:* HARRIS, J. E. & MORRIS, P. E. (eds.) *Everyday memory, actions and absent-mindedness.* London, UK: Academic Press.
HICKOX, A. & SUNDERLAND, A. 1992. Questionnaire and checklist approaches to assessment of everyday memory problems. *In:* CRAWFORD, J. R., PARKER, D. M. & MCKINLAY, W. W. (eds.) *A handbook of neuropsychological assessment.* Hove, England: Lawrence Erlbaum Associates.
HOLTGRAVES, T. M., ROSS, S. J., WEYWADT, C. R. & HAN, T. L. 2007. Perceiving artificial social agents. *Computers in Human Behavior,* 23**,** 2163-2174.

HONE, K. & GRAHAM, R. 2000. Towards a tool for the subjective assessment of speech system interfaces (SASSI). *Natural Language Engineering,* 6**,** 287-305.

HOWE, M. L. & COURAGE, M. L. 2004. Demystifying the beginnings of memory. *Developmental Review,* 24**,** 1-5.

HUMPHREYS, B. L. & LINDBERG, D. A. B. 1993. THE UMLS(R) PROJECT - MAKING THE CONCEPTUAL CONNECTION BETWEEN USERS AND THE INFORMATION THEY NEED. *Bulletin of the Medical Library Association,* 81**,** 170-177.

JANNINK, J. & WIEDERHOLD, G. Ontology maintenance with an algebraic methodology: A case study. AAAI workshop on Ontology Management, 1999.

JONES, G. M. M. & VAN DER EERDEN, W. J. 2008. Designing care environments for persons with Alzheimer's Disease: visuoperceptual considerations. *Clinical Gerontology,* 18**,** 13-37.

JURAFSKY, D. & MARTIN, J. H. 2009. *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition,* Upper Saddle River, N.J., Pearson/Prentice Hall ; London : Pearson Education.

KHAN, R. F. 2011. *The effect of embodied agents within the user interface [electronic resource]* [Online]. University of Manchester. Available: http://www.manchester.ac.uk/escholar/uk-ac-man-scw:125555 [Accessed].

KIM, Y., WEI, Q., XU, B., KO, Y. & ILIEVA, V. 2007. MathGirls: Toward Developing Girls' Positive Attitude and Self-Efficacy through Pedagogical Agents. *Artificial Intelligence in Education,* 158**,** 119-126.

KLUWER, T. 'I Like Your Shirt', - dialogue acts for enabling social talk in conversational agents. Intelligent Virtual Agents, 2011. Springer Lecture Notes in Computer Science, 14-27.

KNIGHT, K. Building a large ontology for machine translation. ARPA Human Language Technology Workshop, 1993 Princeton, USA.

LATHAM, A., CROCKETT, K. & BANDAR, Z. 2010. A conversational expert system supporting bullying and harassment policies. *In:* GROUP, I. S. (ed.). Manchester Metropolitan University.

LATHAM, A. M. 2011. *Personalising Learning with Dynamic Prediction and Adaptation to Learning Styles in a Conversational Intelligent Tutoring System.* PhD Doctoral, Manchester Metropolitan University.

LI, Y., BANDAR, Z. A. & MCLEAN, D. 2003. An approach for measuring semantic similarity between words using multiple information sources. *IEEE Transactions on Knowledge and Data Engineering,* 15**,** 871-882.

LOFTUS, E. F. 1996. *Eyewitness testimony,* Cambridge, MA, USA, Harvard University Press.

LUCAS, R. E. & DYRENFORTH, P. S. 2006. Does the existence of social relationships matter for subjective well-being? *In:* VOHS, K. D. & FINKEL, E. J. (eds.) *Self and Relationships: Connecting Intrapersonal and Interpersonal Processes.* New York, USA: Guildford Press.

MAGEE, J. J. 1994. Using themes from mystical traditions to enhance self-acceptance in life review groups. *Religious Gerontology,* 9**,** 63-72.

MALDE, S. 1988. GUIDED AUTOBIOGRAPHY - A COUNSELING TOOL FOR OLDER ADULTS. *Journal of Counseling and Development,* 66**,** 290-293.

MANN, W. D. & THOMPSON, S. A. 1988. Rhetorical Structure Theory: Towards a functional theory of text organization. *Text,* 8**,** 243-281.
MAULDIN, M. L. 1994. CHATTERBOTS, TINYMUDS, AND THE TURING TEST ENTERING THE LOEBNER-PRIZE COMPETITION. *Proceedings of the Twelfth National Conference on Artificial Intelligence, Vols 1 and 2***,** 16-21.
McKIM, C.A. 2017. The value of mixed methods research: A mixed method study. *Journal of Mixed Methods Research.* 11(2), 202-222
MICHIE, D. & SAMMUT, C. 2000. Infochat Draft Scripter's Manual. Manchester, England.
MICHIE, D., SAMMUT, C., WHEELER, R. & MASON, D. 2006. Museum guide AI system, Australia.
MILLER, G. A. 1995. WORDNET - A LEXICAL DATABASE FOR ENGLISH. *Communications of the Acm,* 38**,** 39-41.
MILWARD, D. & BEVERIDGE, M. 2003. *Ontology-based dialogue systems* [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.136.5607 [Accessed 23 February 2013].
MONGE, P. R., BACHMAN, S. G., DILLARD, J. P. & EISENBERG, E. M. 1982. Communicator competence in the workplace: Model testing and scale development. *Communication Yearbook.* Beverly Hills, California, USA: Sage.
MORSE, A. 2010. Improving Dementia Services in England: An Interim Report. National Audit Office.
MORSE, J.M. & CHUNG, S.E. 2003. Toward holism: The significance of methodological pluralism. *International Journal of Qualitative Methods.* 2(3), 1-12.
MUNN, P. & DREVER, E. 2004. *Using questionnaires in small-scale research*, SCRE.
NAVIGLI, R., VELARDI, P., CUCHIARELLI, A. & NERI, F. Quantitative and Qualitative Evaluation of the OntoLearn Ontology Learning System. 20th International conference on Computational Linguistics., 2004 Switzerland.
NEIMEYER, R. A. 2012. Techniques of grief therapy: Creative practices for counseling the bereaved. New York, USA and London, England: Routledge.
NIELSEN, J. Usability engineering at a discount. Third international conference on human-computer interaction, 1989. New York, NY, USA: Elsevier Science Inc, 394-401.
NIELSEN, J. 2012. How many test users in a usability study. Nielsen Norman Group.
NOY, N. F. & MCGUINNESS, D. L. 2011. *Ontology development 101: A guide to creating your first ontology* [Online]. Stanford University, USA. Available: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html [Accessed 21 February 2013].
NYULAS, C., O'CONNOR, M. & TU, S. Datamaster - a plug-in for importing schemas and data from relational databases into Protégé. 10th International Protégé Conference, 2007 Budapest.
O'SHEA, J. D. 2010. *A framework for applying short text semantic similarity in Goal-Oriented Conversational Agents.* PhD, Manchester Metropolitan University.
O'SHEA, J. D., BANDAR, Z. & CROCKETT, K. 2011. Systems engineering and conversational agents. *In:* TOLK, A. & JAIN, L. C. (eds.) *Intelligence-Based Systems Engineering.* Springer.

O'SHEA, K., CROCKETT, K. & BANDAR, Z. 2010. Application of a Semantic-Based Conversational Agent to Student Debt Management. 2010 IEEE World Congress on Computational Intelligence, 2010. Jul 18-23 2010 Barcelona, SPAIN.

OPPENHEIM, A. N. 1992. *Questionnaire design, interviewing and attitude measurement,* London, Pinter.

OSSHER, L., FLEGAL, K. E. & LUSTIG, C. 2013. Everyday memory errors in older adults. *Aging Neuropsychology and Cognition,* 20**,** 220-242.

PAPATHEODOROU, C., VASSILIOU, A. & SIMON, B. Discovery of ontologies for learning resources using word-based clustering. ED-MEDIA 2002. AACE, 2002 Denver, USA.

PAVOT, W. & DIENER, E. 1993. Review of the satisfaction with Life Scale. *Psychological Assessment,* 5**,** 164-172.

PENN, B. 1994. Using patient biography to promote holistic care. *Nursing times,* 90**,** 35-6.

POLLINA, L. K., GREENE, A. L., TUNICK, R. H. & PUCKETT, J. M. 1992. DIMENSIONS OF EVERYDAY MEMORY IN YOUNG ADULTHOOD. *British Journal of Psychology,* 83**,** 305-321.

PRABHAKAR, J. & HUDSON, J.A. 2018. Past is prologue: The role of memory retrieval in young children's episodic prospection. *Journal of .Experimental Child Psychology.* August 2018.

PRICE, C. & SPACKMAN, K. 2000. SNOMED clinical terms. *British Journal of Healthcare Computing and Information Management,* 17**,** 27-31.

PRINCE, M., GUERCHET, M. & PRINA, M. 2013. Policy brief for Heads of Government. The Global Impact of Dementia 2013-2050. *Alzheimer's Disease International*.

RANK, O. 1924. Das Trauma der Geburt und seine Bedeutung für die Psychoanalyse *Internationale Psychoanalytische Bibliotek*.

RICKEL, J., GANESHAN, R., RICH, C., SIDNER, C. L. & LESH, N. Task-oriented tutorial dialogue: Issues and agents. *In:* ROSE, C. P. F., REVA, ed. AAAI Fall Symposium, 2000 North Falmouth, Massachusettes, USA. AAAI Press, 52-57.

RIGAU, G., RODRIGUEZ, H. & AGIRRE, E. Building Accurate Semantic Taxonomies from Monolingual MRDs. 17th International Conference on Computational Linguistics COLING-ACL'98., 1998 Montreal, Canada.

ROBINSON, S., ROQUE, A. & TRAUM, D. Dialogues in context: An objective User-Oriented Evaluation Approach for Virtual Human Dialogue. 7th International Conference on Language Resources and Evaluation (LREC), 2010 Valletta, Malta.

ROYLE, J. & LINCOLN, N. B. 2008. The Everyday Memory Questionnaire – revised: Development of a 13-item scale. *Disability and Rehabilitation,* 30**,** 114-121.

RUBIN, D. C. 1995. *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes,* New York, USA, Oxford University Press.

RUBIN, D. C., WETZLER, S. E. & NEBES, R. D. 1986. Autobiographical memory across the lifespan. *Autobiographical Memory***,** 202-221.

RYAN, R. M. & DECI, E. L. 2001. On happiness and human potentials: A review of research on hedonic and eudaimonic well-being. *Annual Review of Psychology,* 52**,** 141-166.

SAMMUT, C. 2001. Managing Context in a Conversational Agent. Linkoping Electronic Articles in Computer and Information Science, 3:7 [Online]. http://www.ep.liu.se/ea/cis/2001/007/

SIEGEL, S. & CASTELLAN, N. J. 1988. *Nonparametric statistics for the Behavioral Sciences*, McGraw-Hill.

SILVERVARG, A. & JONSSON, A. Subjective and Objective evaluation of conversational agents in learning environments for young teenagers. Association for the Advancement of Artificial Intelligence, 2011.

SITEPAL. 2012. *What does your site say?* [Online]. http://www.sitepal.com. [Accessed 7 May 2012 2012].

SPROULL, L., SUBRAMANI, M., KIESLER, S., WALKER, J. H. & WATERS, K. 1996. When the interface is a face. *Human-Computer Interaction,* 11**,** 97-124.

STOJANOVIC, L., STOJANOVIC, N. & VOLZ, R. Migrating data-intensive Web Sites into the Semantic Web. 17th ACM symposium on applied computing (SAC), 2002. ACM Press, 1100-1107.

SUNDERLAND, A., BEECH, J. R. & SHEEHAN, E. 1996a. Current orientations test: A study of speed and consistency of retrieval of recent everyday memories by old and young subjects. *Applied Cognitive Psychology,* 10**,** 337-347.

SUNDERLAND, A., HARRIS, J. E. & BADDELEY, A. D. 1983. DO LABORATORY TESTS PREDICT EVERYDAY MEMORY - A NEUROPSYCHOLOGICAL STUDY. *Journal of Verbal Learning and Verbal Behavior,* 22**,** 341-357.

SUNDERLAND, A., HARRIS, J. E. & GLEAVE, J. 1984. MEMORY FAILURES IN EVERYDAY LIFE FOLLOWING SEVERE HEAD-INJURY. *Journal of Clinical Neuropsychology,* 6**,** 127-142.

SUNDERLAND, A., STEWART, F. M. & SLUMAN, S. M. 1996b. Adaptation to cognitive deficit? An exploration of apparent dissociations between everyday memory and test performance late after stroke. *British Journal of Clinical Psychology,* 35**,** 463-476.

SUNDERLAND, A., WATTS, K., BADDELEY, A. D. & HARRIS, J. E. 1986. SUBJECTIVE MEMORY ASSESSMENT AND TEST-PERFORMANCE IN ELDERLY ADULTS. *Journals of Gerontology,* 41**,** 376-384.

SURYANTO, H. & COMPTON, P. Discovery of Ontologies from Knowledge Bases. *In:* SHAVLIK, T. G. M. M. J., ed. first International Conference on Knowledge Bases., 2001 Victoria, British Columbia Canada. The Association of Computing Machinery, New York, USA., 171-178.

THORNDIKE, E. L. 1905. *The elements of psychology,* New York, USA, Seiler.

TINSON, D. J. & LINCOLN, N. B. 1987. Subjective memory impairment after stroke. *International disability studies,* 9**,** 6-9.

TOSELAND, R. W. 2005. What's different about working with older people in groups? *Gerontological Social Work,* 44**,** 5-23.

TOURANGEAU, R., RIPS, L. J. & RASINSKI, K. A. 2000. *The psychology of survey response,* Cambridge, U.K., Cambridge University Press.

TULVING, E. 1983. *Elements of Episodic Memory,* Oxford, Clarendon Press.

TULVING, E. 2002. Episodic Memory: From mind to brain. *Annual Review of Psychology,* 53**,** 1-25.

TURING, A. M. 1950. Computing Machinery and Intelligence. *Mind***,** 433-460.

VAN ASSEM, M., MENKEN, M. R., SCHREIBER, G., WIELEMAKER, J. & WIELINGA, B. J. A method for converting thesauri to RDF/OWL. International Semantic Web Conference., 2004. 17-31.

VAN DE WATERING, M, 2018. The impact of Computer Technology on the Elderly. Vrije Universiteit, Amsterdam.

VELARDI, P., NAVIGLI, R., CUCHIARELLI, A. & NERI, F. 2005. Evaluation of ontolearn, a methodology for automatic population of domain ontologies. *In:* BUITELAAR, P. C., P. ;MAGNINI,B (ed.) *Ontology Learning from Text:Methods, Applications and Evaluation.*: IOS Press.

VOLZ, R., OBERLE, D., STAAB, S. & STUDER, R. 2003. OntoLiFT Prototype. *IST Project 2001-33052 WonderWeb Deliverable 11.*

WAGNER, N., HASSANEIN, K. & HEAD, M. 2010. Computer use by older adults: A multi-disciplinary review. *Computers in Human Behavior,* 26**,** 870-882.

WALKER, J. H., SPROULL, L. & SUBRAMANI, R. 1994. Using a human face in an interface. *Human Factors in Computing Systems, Chi '94 Conference Proceedings - Celebrating Interdependence***,** 85-91.

WALKER, M. A., LITMAN, D. J., KAMM, C. A. & ABELLA, A. 1997. PARADISE: A framework for evaluating spoken dialogue agents. *35th Annual Meeting of the Association for Computational Linguistics and the 8th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference***,** 271-280.

WALLACE, R. 2016. *Intelligence Markup Language - An introduction* [Online]. https://pandorabots.com/docs/aiml/aiml-basics.html. [Accessed 18 May 2018 2018].

WALLACE, R. S. 2009. The anatomy of A.L.I.C.E. *Parsing the Turing Test Part III.* Berlin: Springer.

WALTER, S. & MEIER, B. 2014. How important is importance for prospective memory? A review. *Front Psychol,* 5**,** 657.

WATSON, D., CLARK, L. A. & TELLEGEN, A. 1988. DEVELOPMENT AND VALIDATION OF BRIEF MEASURES OF POSITIVE AND NEGATIVE AFFECT - THE PANAS SCALES. *Journal of Personality and Social Psychology,* 54**,** 1063-1070.

WEIZENBAUM, J. 1966. ELIZA - A computer program for the study of natural language communication between man and machine. *Communications of the Association for Computing Machinery,* 9**,** 36-45.

WILCOX, B. 2011. *Beyond Facade - Pattern Matching for Natural Language Applications* [Online]. http://chatscript.sourceforge.net/Documentation/Paper-%20Pattern_Matching_for_Natural_Language_Applications.pdf. [Accessed 31 August 2013 2013].

WILSON, B., BADDELEY, A. & COCKBURN, J. 1986. THE RIVERMEAD BEHAVIORAL MEMORY TEST - REPORT OF THE PILOT-STUDY. *Bulletin of the British Psychological Society,* 39**,** A56-A56.

ZHAO, H. 2000. Universal Usability Web Design Guidelines for the Elderly (Age 65 and Older). University of Maryland, USA.

ZIGMOND, A. S. & SNAITH, R. P. 1983. THE HOSPITAL ANXIETY AND DEPRESSION SCALE. *Acta Psychiatrica Scandinavica,* 67**,** 361-370.

Appendix 1 Ontology class examples

Classes used within the ontology

Classes used within the reminiscence ontology for themed knowledge

Appendix 2 Participant information Initial Experiments

Study information

We would like to invite you to take part in our research study. Before you decide we would like you to understand why the research is being done and what it would involve for you. The study will be research into the use of a conversational agent as a memory aid for people suffering from normal ageing memory loss.

A conversational agent is a means of communicating between a human and a computer using typed input. The conversational agent will interact with you by using reminiscence. It will allow you to revisit past events both personally and nationally.

In the beginning, we shall record interviews with you about your life story. From this, keywords will be collected and stored on the agent. Once a knowledge base is built up, questionnaires will be collected anonymously, allowing you to be very honest about your experience with the agent.

If the study is successful, it is believed that the agent will be used as a treatment to delay the effects of ageing memory loss by providing stimulation for the user's memory and allowing lengthy conversations

Ethics statement for participants

Call for participants in an experimental study using a conversational agent to improve ageing memory loss.

We would like to request your participation in a scientific study of a reminiscing conversational agent. For ethical reasons, we are required to ask your permission in advance and let you know what you are agreeing to.

We have provided the answers to the key ethical questions below.

If you require any further information before agreeing to participate, please contact Collette Curry (collette.curry@mmu.ac.uk).

What will you ask me to do?

If you agree, you will be asked to take part in recorded Life Story interviews where key- words will be extracted for use with the agent. You will also be asked a few questions about yourself and about the agent. These are your name, age and the highest level of qualification. You will also be asked to confirm that you are a native English speaker (i.e. Someone for whom English is their first language, spoken since birth). We ask for some personal data because sometimes scientific studies produce surprising results which need to be analysed and this background information could help this analysis.

Is there any risk?

The study does not pose any risk to yourself, nor will it be offensive. The risk involved is equivalent to talking with a friend or family member.

How long will the data be kept for?

The answers to the questions about yourself will be kept for no longer

than is necessary to check for errors or interesting properties of the data. This will be for no longer than 3 months after the first results are published. The ratings you provide will be separated from the personal data and kept permanently. This is because the data can be very useful in long-term studies. Data of this type collected in the 1960s is still widely used today.

Will you publish my personal information?

No personal information will be published.

Appendix 3 Consent form for participants, Initial Experiments

Consent Form for conversational agent reminiscence aid

Thank you for considering taking part in this research. If you have any questions, please ask the researcher before you decide whether to take part.

You will be given a copy of this consent form to keep and refer to at any time. I confirm that I have read and understood the information sheet dated

for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

I understand that my participation is voluntary and that I am free to withdraw at any time without explanation.

I understand that relevant sections of my notes and data collected during the study may be looked at by individuals from the Intelligent Systems Group, where it is relevant to my taking part in the research. I give permission for these individuals to have access to my records.

I understand that if I withdraw from the study, the data collected up to that point will be destroyed.

I agree to take part in the study.

Participant name:

Date:

Researcher name:

Date:

Appendix 4 Participant Questionnaire prior to using the reminiscing CA framework for initial experiments

Questionnaire of reminiscing

Please answer the following questions as fully as possible. The answers will help to produce a reminiscence aid on the computer. There are no right or wrong answers, just write down as many or as few answers as you can.

1. What do you remember most about yourself as a younger person?

2. You may have lived through several decades, what do you recall most about each decade?

3. Please write down keywords that might prompt you to reminisce.

Appendix 5 How to access Betty for participants

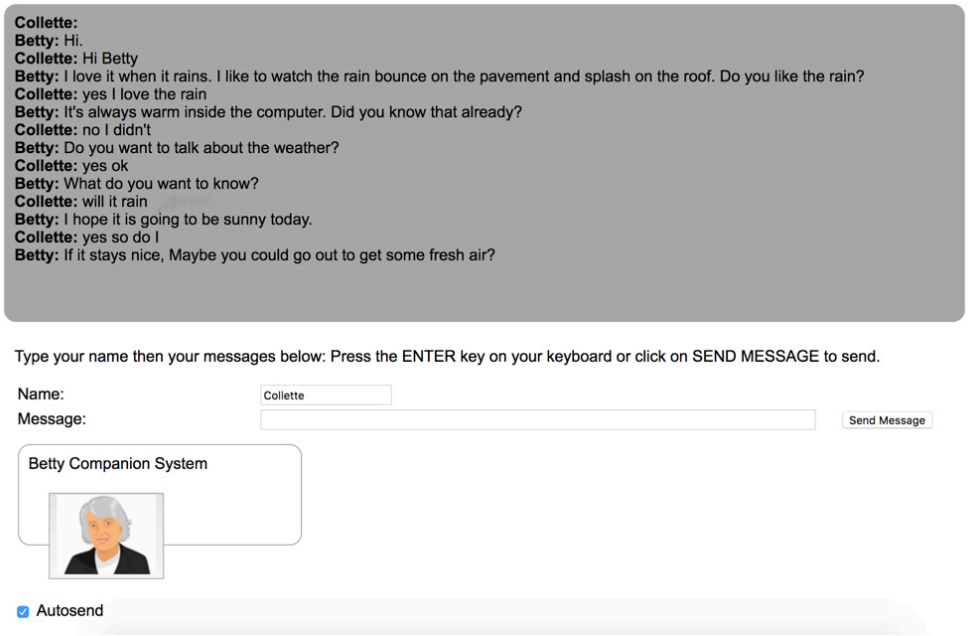
Dear participant,

Thank you for agreeing to talk with Betty for 15 minutes a day for two weeks. To find Betty on your computer follow the instructions below.

Open the web browser that you normally use to look at web pages.
In the address bar at the top of the page type in the following web address:

bit.ly/bettychatgo

A similar screen to that shown below will be opened. Type in your name in the 'Name' box. Then type in your message in the box below your name. Press the 'ENTER' key on your keyboard or click on the 'Send Message' button to send your message to Betty. Once Betty answers, you can carry on sending messages to Betty by typing a message and sending it.



Your conversations with Betty will be saved as a log file anonymously. You can talk about anything that you like. Betty enjoys talking about her school days, the weather and pets most of all.

Any problems, please contact me at my email address **collette.curry@stu.mmu.ac.uk** and I shall help to sort the problems out.

Appendix 6 Everyday Memory Questionnaire 20Q

Here is a set of memory problems which we all experience to some extent in everyday life. Please think back over the PAST WEEK on how often each one has happened to your partner. Tick the circle "same as a year ago" if you think their forgetfulness has remained the same over the past year or "worse" if you think this type of forgetfulness has become more frequent. It may also be possible that their memory has improved therefore tick "better".

Having done this, again think back over the PAST WEEK and write down your estimate of how often each one has happened to your partner. For example, if you think he/she has not lost anything around the house over the past week, put 0 in the square brackets beside item 1; if you think it has happened about once a day then put 7, etc. It is impossible to answer these questions exactly, but please make the best guess you can and do not leave any blanks.

1. **Forgetting where he or she has put something.**
   Losing things around the house.

Same as a year ago □ Worse □ Better □ How often this week [ ]

2. **Forgetting someone's name or calling them by the wrong name.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

3. **Telling someone a story or a piece of news that he/she has already told them.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

4. **Forgetting what day it is.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

5. **Having to go back to check whether he or she has done something that he/she meant to do.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

6. **Forgetting when it was that something happened; for example, whether it was yesterday, or last week.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

7. **Getting confused when trying to explain something.**

Same as a year ago □ Worse □ Better □ How often this week [ ]

8. Forgetting that he/she was told something yesterday or a few days ago.

Same as a year ago □ Worse □ Better □ How often this week [ ]

9. Starting to read a book or a newspaper story without realising that he/she had read it already.

Same as a year ago □ Worse □ Better □ How often this week [ ]

10. Letting him or herself ramble on to talk about irrelevant or unimportant things.

Same as a year ago □ Worse □ Better □ How often this week [ ]

11. Forgetting details of a regular routine-forgetting what to do or when to do it.

Same as a year ago □ Worse □ Better □ How often this week [ ]

12. Having difficulty in picking up a new skill. For example, finding it hard to learn a new game or how to work a new gadget.

Same as a year ago □ Worse □ Better □ How often this week [ ]

13. When talking to someone, unable to find the right word to use.

Same as a year ago □ Worse □ Better □ How often this week [ ]

14. Forgetting to do something he or she meant to do.

Same as a year ago □ Worse □ Better □ How often this week [ ]

15. Forgetting important details of what he/she did, or what happened the day before.

Same as a year ago □ Worse □ Better □ How often this week [ ]

16. Getting lost or turning in the wrong direction.

Same as a year ago □ Worse □ Better □ How often this week [ ]

17. Being unable to follow the thread of a story or television programme.

Same as a year ago □ Worse □ Better □ How often this week [ ]

18. Forgetting to tell somebody something important. Perhaps forgetting to pass on a message or remind someone of something.

Same as a year ago □ Worse □ Better □ How often this week [ ]

19. Failing to recognise the faces of people he/she should know.

Same as a year ago □ Worse □ Better □ How often this week [ ]

20. Getting the details of what someone has told him or her mixed up.

Same as a year ago □ Worse □ Better □ How often this week [ ]

In everyday life, how much of a problem would you say memory and concentration difficulties are to your partner?

(Please tick one of the following)

No problem □
Slight problem □
Moderate problem □
Severe problem □

Appendix 7 Everyday Memory Questionnaire (13)

Instructions
Below are listed some examples of things that happen to people in everyday life. Some of them may happen frequently and some may happen very rarely. We should like to know how often on average you think each one has happened to you over the past month. Write the appropriate letter in the box beside the item.

A. Once or less in the last month.
B. More than once a month but less than once a week.
C. About once a week.
D. More that once a week or less than once a day.
E. Once or more in a day.

1. Having to check whether you have done something that you should have done. ☐
2. Forgetting when it was that something happened; for example, whether it was yesterday or last week. ☐
3. Forgetting that you were told something yesterday or a few days ago, and maybe having to be reminded about it. ☐
4. Starting to read something (a book or an article in a newspaper, or a magazine) without realizing you have already read it before. ☐
5. Finding that a word is 'on the tip of your tongue'. You know what it is but cannot quite find it. ☐
6. Completely forgetting to do things you said you would do, and things you planned to do. ☐
7. Forgetting important details of what you did or what happened to you the day before. ☐
8. When talking to someone, forgetting what you have just said. Maybe saying 'what was I talking about?' ☐
9. When reading a newspaper or magazine, being unable to follow the thread of a story; losing track of what it is about. ☐
10. Forgetting to tell somebody something important, perhaps forgetting to pass on a message or remind someone of something. ☐
11. Getting the details of what someone was told you mixed up and confused. ☐
12. Forgetting where things are normally kept or looking for them in the wrong place. ☐
13. Repeating to someone what you have just told them or asking someone the same question twice. ☐

Appendix 8 HADS Questionnaire

Each item on the HADS questionnaire is scored from 0 for 'positive' responses to 3 for 'negative' responses to the 14 questions.

Hospital Anxiety and Depression Scale (HADS)

Question	Nearly all the time	Very often	Sometimes	Not at all
I feel tense or wound up				
I feel as if I am slowed down				
I still enjoy the things I used to enjoy				
I get a sort of frightened feeling like 'butterflies' in the stomach				
I get a sort of frightened feeling as if something awful is about to happen				
I have lost interest in my appearance				
I can laugh and see the funny side of things				
I feel restless as I have to be on the move				
Worrying thoughts go through my mind				
I look forward with enjoyment to things				
I feel cheerful				
I get sudden feelings of panic				

I can sit at ease and feel relaxed				
I can enjoy a good book or radio or TV programme				

Appendix 9 EMQ-R Questionnaire issued to participants

Below are listed some examples of things that happen to people in everyday life. Some may happen frequently and some may happen very rarely. We would like to know how often on average you think each one has happened to you over the past month. Write the appropriate letter in the box beside the item.

	Once or less in the last month	**A**
	More than once a month but less than once a week	**B**
	About once a week	**C**
	More than once a week or less than once a day	**D**
	Once or more in a day	**E**
1	Having to check whether you have done something that you should have done.	
2	Forgetting when it was that something happened; for example, whether it was yesterday or last week.	
3	Forgetting that you were told something yesterday or a few days ago, and maybe having to be reminded about it.	
4	Starting to read something (a book or an article in a newspaper or a magazine) without realising you have already read it before.	
5	Finding that a word is 'on the tip of your tongue'. You know what it is but cannot quite find it.	
6	Completely forgetting to do things you said you would do, and things you planned to do.	
7	Forgetting important details of what you did or what happened to you the day before.	
8	When talking to someone, forgetting what you have just said. Maybe saying "What was I talking about?".	
9	When reading a newspaper or magazine, being unable to follow the thread of a story; losing track of what it is about.	
10	Forgetting to tell somebody something important, perhaps forgetting to pass on a message or remind someone of something.	
11	Getting the details of what someone has told you mixed up and confused.	
12	Forgetting where things are normally kept or looking for them in the wrong place.	

13	Repeating to someone what you have just told them or asking someone the same question twice.	
	PLEASE CHECK THAT YOU HAVE PUT A LETTER IN EVERY BOX.	

Appendix 10 SPANE Instrument Questionnaire

Scale of Positive and Negative Experience (SPANE)

Please consider what you feel and experience during the next two weeks. Then report how much you experienced each of the following feelings, using the scale below. You will have been given three copies of this form. Please complete this form before talking with Betty, then at day 7 and finally at day 14.

For each item, select your answer, and put a tick in the table below for each item.

BEFORE TALKING	**Very rarely or Never**	**Rarely**	**Sometimes**	**Often**	**Very often or always**
Positive					
Happy					
Good					
Joyful					
Pleasant					
Contented					
Negative					
Sad					
Afraid					
Bad					
Angry					
Unpleasant					

Before talking with Betty

ON DAY SEVEN	Very rarely or Never	Rarely	Sometimes	Often	Very often or always
Positive					
Happy					
Good					
Joyful					
Pleasant					
Contented					
Negative					
Sad					
Afraid					
Bad					
Angry					
Unpleasant					

On day seven of talking with Betty

ON LAST DAY	Very rarely or Never	Rarely	Sometimes	Often	Very often or always
Positive					
Happy					
Good					
Joyful					
Pleasant					
Contented					
Negative					
Sad					
Afraid					
Bad					
Angry					
Unpleasant					

On day fourteen of talking with Betty

Appendix 11 Likert scale 18-part questionnaire

Please place a tick on the chosen answer for each of the following 18 questions.

No.	Question	**Strongly disagree**	**Disagree**	**Don't know**	**Agree**	**Strongly agree**
1	Betty's answers often surprised me					
2	Betty understood what I said					
3	I could fix misunderstandings with Betty if I wanted to					
4	Betty was a good listener					
5	I would like to talk to Betty again					
6	Betty expresses her ideas very clearly					
7	Betty mostly says the right thing at the right time					
8	Betty is easy to talk to					
9	I liked to talk to Betty					
10	I could control the interaction with Betty					
11	It was easy to understand how to talk so that Betty could understand					
12	It felt natural to talk to Betty					
13	Sometimes I lost track of the conversation					

14	It was frustrating to talk to Betty					
15	It was hard to know what to talk about with Betty					
16	Betty often repeated herself					
17	Sometimes I wondered if I used the right word					
18	I always knew what I could say to Betty					

[Based on a Likert Scale questionnaire carried out by Silvervarg & Jonsson (Silvervarg and Jonsson, 2011)]

Appendix 12 Dialogue action codes and CA response values for future objective analysis

Code	Description
D	General dialogue functions, e.g. Greeting, Closing, Politeness
H	Hazing, Testing
Q	Information request
R	Requests, in our case for general information
A	Answer to CA utterances
S	Statements
I	CA initiatives
REP	Repeated CA utterances
G	Incomprehensible, typically empty utterances or nonsense

Dialogue action codes

Code	Value
3	Appropriate response
2	Partially appropriate
RR	Request Repair
NR1	No response (Inappropriate)
NR3	No response (Appropriate)
1	Inappropriate response

CA response values coding scheme

Sources: (Silvervarg and Jonsson, 2011)

Appendix 13 Keywords

Example list of relevant keywords

Childhood in 20's, 30's, 40's, 50's, 60's, 70's

Toys and games

Street play

Cobbled streets and pitch (tar)

The knocker upper

Clothes and shoes

Sweet shops

Penny mix

Sparse, drab furniture/decor

Flea pit cinemas

Haporth of chips

School

Nit nurse

Purple ringworm paint

Church

Deep snowy winters

Fairgrounds

Bommy night

Open fires

Christmas

Whit walks

Wakes week

Leaving school

First job

Boyfriends

Girlfriends

The pub

Dance halls

Belle Vue Zoo

Blackpool

Corner shops

Marriage

First home

Parenthood

Music

Films/film stars

Fads/fashions

Newspapers/magazines

Jackie, Bunty, Beano

War years

Rationing

Appendix 14 Conversational Logs

Example conversational logs keywords and themes collected

Horses and carts, trams, steam trains, old cars (Wolseley's, Crossley's, Austin Mini cars)

Trams the rod clattering down from the overhead cables and it having to stop while the conductor hooked it back on with a big pole. The noise of it made everyone jump.

Hustle and bustle, sounds, sights and smells now long gone

National health specs
School milk

All over wash at the sink, stood on a towel using a block - cube of green fairy soap or of red carbolic soap that smelled like a burning pan handle.

String vests

Brylcream

Shavers and blades

Old spice aftershave

Tin bath by the open fire once a week

Outside bog-cludgie-netty

Back to back houses with small yards, back entry, scraping

of metal dustbins

Pea soupers - Foggy days and nights

Flagstones

Donkey stones

Ragbone man

Roaches, silverfish and mice

Crofts - bommies

Milkman

Rent man

The club man

Hiding, pretending not to be in when anyone wanting payment knocked on, when Mams were a bit short that week

Going to pay the papers

Taking the bottles back

The mineral man or pop man

Chippy tea usually on fri or sat

50's first tellies, first phones, fridge, washing machine

60's fashions, hairstyle, music, dance crazes, hair do's

70's party food; food on cocktail sticks stuck into a foil covered ta-ter, other popular food

Party7 cans of beer/lager

Popular telly progs. Few channels. T V times, Look In,
wheel tappers,

THE WIRELESS

GRAMAPHONES

LARD

SCRATCHINGS IN THE CHIPPY;

NEWSPAPER WRAPPING FOR CHIPS

NEWSPAPER SQUARES IN THE CLUDGIE INSTEAD OF BOGPAPER
a loop of string was pushed through the pieces and hung on a nail next to the toilet.

BOX BROWNIE CAMERAS

MACASSER OIL - BRYL CREAM
CUT THROAT RAZORS-STROP
CHILDRENS HOUR (50'S RADIO)

WATCH WITH MOTHER (50'S, 60'S ?TELLY)

ILLNESSES-EPIDEMICS; SCARLET FEVER, INFLUENZA, TB, POLIO, PNEUMONIA, WHOOPING COUGH, CHICKEN POX, MEASLES, DIABETES (BEFORE INSULIN 1924)

NO NHS LOW WAGES

TIMES WERE VERY HARD

HARD TO KEEP WARM DURING THE HARSH WINTERS

WORKHOUSE-POORHOUSE. LIKE PRISONS

HORSE DRAWN TRAMS, FIRE ENGINES, MILK FLOATS, RAGBONE MEN OR PUSHING FLAT HANDCARTS

POVERTY, MALNUTRITION, OVERCROWDING

CHILD LABOUR

SHORT EDUCATION IF ANY

PLAYGROUND - STREET GAMES; SKIPPING AND TWO BALL WHILE SINGING OR CHANTING;

salt , pepper, vinegar, mustard.

Ring a ring of roses,

hot cross buns, hot cross buns, one a penny two a penny hot cross buns;

girls would stand in a tight semi-circle with toes of one shoe touching against everyone elses while the allotted dipper would touch each shoe as she was chanting, one word at a time dip, dip, dip, my blue ship, sailing on the water, like a cup and saucer

out goes you (to decide who was on during hide and seek etc.)

Here comes the muffin man, the muffin man, the muffin man; here comes the muffin man who lives on Drury lane.

The farmers in his den, the farmers in his den, eee I endy oh the farmers in the den; the farmer wants a wife..

What's your name; Mary Jane, where dya live; down the grid, what number; cucumber.

The big ship sails on the alley alley o

Eeeny meany miny mo.

In and out the fairy bluebells.

THE FLICKS,

PICTURES,

CINEMA AND THEATRE IN EVERY TOWN

POPULAR ALCOHOLIC BEVERAGES;

MACKIES,

PALE ALE,

MILK STOUT,

Babycham,

CHERRY B,

HOT CHESTNUT SELLER

TRAMPS-URCHINS-ORPHANS?

ORGAN GRINDER AND MONKEY?

THE PARLOUR OR BEST ROOM FOR GUESTS-IMPORTANT PEOPLE, it was out of bounds to kids.

Gas lamps-street lights

THE WAR YEARS first, as seen by a child

FRUGAL FOOD;

cheese and bacon,

cow heel stew,

corned beef hash,

ribs & cabbage,

potato hash,

bread - toast and dripping,

mutton,

bubble and squeak,

black pudding,

kidney,

liver,

brawn,

tongue,

heart,

brains,

sheep's head,

rabbit,

tripe, chitterlings, faggots

pea soup and bacon shank (pestle)

Fish and chips hap 'Orth o chips

Toast was made on a toasting fork at the open fire.
Chestnuts on the shovel in fire

Public baths once a week for sixpence a bath
Swimming baths